RAPID MIXING OF HYPERGRAPH INDEPENDENT SETS

JONATHAN HERMON[‡], ALLAN SLY[†,*], AND YUMENG ZHANG[*]

Abstract. We prove that the the mixing time of the Glauber dynamics for sampling independent sets on n-vertex k-uniform hypergraphs is $O(n \log n)$ when the maximum degree Δ satisfies $\Delta \leqslant c2^{k/2}$, improving on the previous bound [2] of $\Delta \leqslant k-2$. This result brings the algorithmic bound to within a constant factor of the hardness bound of [1] which showed that it is NP-hard to approximately count independent sets on hypergraphs when $\Delta \geqslant 5 \cdot 2^{k/2}$.

1. Introduction

We consider the mixing time of the Glauber dynamics for sampling from uniform independent sets on a k-uniform hypergraph (i.e., all hyperedges are of size k). In doing so we extend the region where there is a fully polynomial-time randomized approximation scheme (FPRAS) for approximately counting independent sets, reducing an exponential multiplicative gap to a constant factor.

In the case of graphs the question of approximately counting and sampling independent sets is already well understood. In a breakthrough paper, Weitz [15] constructed an algorithm which approximately counts independent sets on 5-regular graphs by constructing a tree of self-avoiding walks to calculate marginals of the distribution. These can be approximated efficiently because of decay of correlation giving rise to a fully polynomial-time approximation scheme (FPTAS) for the problem. This was shown to be tight [13] via a construction based on random bipartite graphs, proving that it is NP-hard to approximately count independent sets on 6-regular graphs. The key difference between 5 and 6 is that on the infinite 5-regular tree, there is exponential decay of correlation of random independent sets while long range correlations are possible on the 6-regular tree.

In terms of statistical physics the difference is that there is an unique Gibbs measure on the Δ-regular tree for $\Delta \leqslant 5$ but the existence of multiple Gibbs measures when $\Delta \geqslant 6$. This paradigm extends more broadly to other spin systems such as the hardcore model (a model of weighted independent sets) and the anti-ferromagnetic Ising model. In both cases a similar construction to [15, 13] shows that it is NP-hard whenever these models have non-uniqueness [14] and Weitz's algorithm gives an FPTAS [12] in the uniqueness case except for certain critical boundary cases. Together with work of Jerrum and Sinclair [5] the problem of approximately counting in 2-spin systems on regular graphs is essentially complete.

For hypergraphs, however, even in two spin systems the question remains wide open. A hypergraph $H = (V, F)$ consists of a vertex set V and a collection F of vertex subsets, called the hyperedges. An independent set of H is a set $I \subseteq V$ such that no hyperedge $a \in F$ is a subset of I. The natural analogy with graphs would predict that the threshold for

Princeton University[†], University of Cambridge, Cambridge, UK, Financial support by the EPSRC grant EP/L018896/1[‡] and University of California, Berkeley[*]

E-mail addresses: jonathan.hermon@statslab.cam.ac.uk, sly@berkeley.edu, ymzhang@berkeley.edu.

Key words and phrases. Mixing time, hypergraph independent sets, approximate counting.

approximate counting should correspond to the uniqueness threshold for the Δ-regular tree which corresponds to $\Delta = (1+o(1))\frac{1}{2k}2^k$. This turns out to be false and in fact [1] showed that it is NP-hard to approximately count independent sets when $\Delta \geqslant 5 \cdot 2^{k/2}$. What breaks down is that Weitz's argument requires not just exponential decay of correlation but also a stronger notion known as strong spatial mixing (SSM) which fails to hold when $\Delta \geqslant 6$ for all $k \geqslant 2$ [1].

Despite the lack of SSM, Bezáková, Galanis, Goldberg, Guo, Štefankovič [1] were able to give a modified analysis of the Weitz's tree of self avoiding walks algorithm and gave a deterministic FPTAS for approximating the number of independent sets when $200 \leqslant \Delta \leqslant k$. In this paper we study the Glauber dynamics where previously using path coupling Bordewich, Dyer and Karpinski [3, 2] showed that the mixing time is $O(n \log n)$ when $\Delta \leqslant k-2$ (where throughout n is used to denote the number of vertices). These bounds, while holding for larger Δ than the graph case, still fall far short of $5 \cdot 2^{k/2}$, the hardness bound. Our main result gives an improved analysis of the Glauber dynamics narrowing the computational gap to a multiplicative constant. In the case of *linear* hypergraphs, those in which no hyperedges share more than one vertex, much stronger results are possible.

Theorem 1.1. *There exists an absolute constant $c > 0$ such that for every n-vertex hypergraph G with edge size k and maximal degree Δ, the Glauber dynamics mixes in $O(n \log n)$ time if the graph satisfies one of the following conditions:*
1. $\Delta \leqslant c2^{k/2}$.
2. $\Delta \leqslant c2^k/k^2$ and G is linear.

We expect our approach to hold when the sizes of hyperedges are at least k but for the sake of simplicity of the proof we restrict our attention to the case of constant hyperedge size. When the hypergraph is linear we achieve a much stronger bound of $\Delta \leqslant c2^k/k^2$, close to the uniqueness threshold of $\Delta = (1+o(1))\frac{1}{2k}2^k$. This suggests the possibility that it may only be the presence of hyperedges with large overlaps that is responsible for the discrepancy with the tree uniqueness threshold. Indeed, in the hardness construction of [1] pairs of hyperedges have order k vertices in common.

Our mixing time proof directly translates into an algorithm for approximately counting independent sets.

Corollary 1.2. *There is an FPRAS for counting the number of hypergraph independent sets for all hypergraphs with maximal degree Δ and edge size k satisfying the conditions of Theorem 1.1.*

Closely around the time a first version of this manuscript was posted on arXiv, two related results were posted by different groups of authors: Moitra [10] gave a new FPTAS up to $\Delta \leqslant 2^{k/20}$; Heng, Jerrum and Liu [4] gave an *exact* sampling algorithm that has $O(n)$-average running time when $\Delta \leqslant \frac{1}{\sqrt{6ek}}2^{k/2}$ and the minimum intersection $s \geqslant O(\log \Delta + \log k)$ between any pair of hyperedges. Both results were inspired by the recent breakthroughs on algorithmic Lovász Local Lemma [11], but took significantly different approaches beyond that. While not giving as sharp results as ours in the case of hypergraph independent sets (i.e., monotone CNF formulas), the two algorithms apply to general CNF formulas. It also worth noticing that while our algorithm works better when neighbouring hyperedges have small intersections, the algorithm in [4] works best when the intersections are large.

We also consider the case of a random regular hypergraph which is of course locally treelike. Let $\mathcal{H}(n, d, k)$ be the uniform measure over the set of hypergraphs with n vertices, degree d and edge size k. In this case we are able to prove fast mixing for d growing as $c2^k/k$ which is the same asymptotic as the uniqueness threshold.

Theorem 1.3. *There exists an absolute constant $c > 0$ such that if H is a random hypergraph sampled from $\mathcal{H}(n, d, k)$ with $d \leqslant c2^k/k$, then with high probability (over the choice of H), the Glauber dynamics mixes in $O(n \log n)$ time.*

The only property of random regular hypergraphs used in the proof of Theorem 1.3 is that there exists some constant $N \geqslant 1$ such that each ball of radius $R = R(\Delta, k) \equiv \lceil 3e\Delta k^2 \rceil$ contains at most N cycles. Indeed, if $H \sim \mathcal{H}(n, d, k)$ then this holds for $N = 1$ with high probability for every fixed radius (cf. [8, Lemma 2.1]).

1.1. **Proof Outline.** As noted above two key methods for approximate counting, tree approximations and path coupling break down far from the computational threshold. Tree approximations rely on a strong notion of decay of correlation, strong spatial mixing, which as noted above breaks down even for constant sized Δ and the work Bezáková et al. [1] to extend to Δ growing linearly in k required a very detailed analysis. Similarly, for the Glauber dynamics, path coupling also breaks down for linear sized Δ [2].

It is useful to consider the reasons for the limitations of path coupling. Disagreements can only be propagated when there is a hyperedge with $k - 1$ ones. However, such hyperedges should be very rare. Indeed, we show that in equilibrium the probability that a certain hyperedge a has $k - 1$ ones is at most $(k + 1)2^{-k}$. Thus when Δ is small, most vertices will be far from all such hyperedges which morally should give a contraction in path coupling. However, in the standard approach of path coupling we must make a worst case assumption of the neighbourhood of a disagreement.

Our approach is to consider the geometric structure of bad regions in space time $V \times \mathbb{R}_+$. In Section 3 we give a simpler version of the proof which loses only a polynomial factor in the bound on Δ, yet highlights the key ideas that will be used later. We bound the bad space time regions via a percolation argument showing that if coupling fails, then some disagreement at time 0 must propagate to the present time, which corresponds to a vertical crossing. This geometric approach is similar to the approach of Information Percolation used to prove cutoff for the Ising model [9] and avoids the need to assume worst case neighbourhoods.

In Section 3 we control the propagation of disagreements by discretizing the time-line into blocks of length k and considering some (fairly coarse) necessary conditions for the creation of new disagreements during the span of an entire block of time. In particular, in order for a new disagreement to be created at $a \in F$ during a given time interval I, there must be some $t \in I$ at which the configuration on the vertices of a has $k - 1$ ones. This allows us to exploit the independence between different time blocks. The proof for the sharp result is given in Section 4 and Section 5 where a more refined analysis is carried out. The main additional tool is to find an efficient scheme for controlling the propagation of disagreements via an auxiliary continuous time process so that we again can exploit independence, as well as certain positive correlations associated with that process.

Finally, in Section 6 we present the proof of Corollary 1.2 via a standard reduction from sampling to counting.

2. Preliminaries

2.1. **Definition of model.** In what follows it will be convenient to treat vertices and hyperedges in a uniform manner, for which reason we consider the bipartite graph representation $G = (V, F, E)$ of a hypergraph $H = (V, F)$, where V is the set of vertices, F is the set of hyperedges, and $E = \{(v, a) : v \in a \in F\}$ (i.e., we connect vertex $v \in V$ to hyperedge $a \in F$ if and only if v appears in hyperedge a). Let $n \equiv |V|$ denote the number of vertices in G. For each $v \in V$, we will denote by ∂v the neighbours of v in G, which is a subset of F and for each $a \in F$ define ∂a similarly. Under this notation, the degree of a vertex v equals $|\partial v|$ while the size of a hyperedge a equals $|\partial a|$.

An *independent set* of hypergraph G can be encoded as a configuration $\underline{\sigma} \in \{0,1\}^V$ satisfying that for every $a \in F$, there exists $v \in \partial a$ such that $\sigma(v) \neq 1$. We denote by $\Omega \equiv \Omega(G) \subset 2^V$ the set of all such configurations and consider the uniform measure over Ω given by

$$\pi(\underline{\sigma}) = \frac{1}{Z(G)} \mathbf{1}\{\underline{\sigma} \text{ is an independent set of } G\},$$

where the normalizing constant $Z \equiv Z(G) \equiv |\Omega(G)|$ counts the number of hypergraph independent sets on G.

The (discrete-time) *Glauber dynamics* on the set of independent sets is the Markov chain $(W_t)_{t \geqslant 0}$ with state space Ω defined as following: For each configuration $\underline{\sigma} \in \Omega$, vertex $v \in V$ and binary variable $x \in \{0,1\}$, let $\underline{\sigma}^{v,x}$ be the configuration that equals x at vertex v and agrees with $\underline{\sigma}$ elsewhere. Suppose that the Markov chain is at state $W_t = \underline{\sigma}$ at time t. The state at time $t+1$ is then determined by uniformly selecting a vertex $v \in V$ and performing the following update procedure:

1. With probability $1/2$, set $W_{t+1} = \underline{\sigma}^{v,0}$.
2. With the rest probability $1/2$, set $W_{t+1} = \underline{\sigma}^{v,1}$ if $\underline{\sigma}^{v,1} \in \Omega$ and $W_{t+1} = \underline{\sigma}$ otherwise. (If the latter case happens then $W_{t+1}(v) = \sigma(v) = 0$.)

This Markov chain is easily shown to be ergodic with stationary distribution π. Its (total variation) mixing time, denoted by t_{mix}, is defined to be

$$t_{\text{mix}}(\epsilon) \equiv \inf\{t : \max_{\underline{\sigma} \in \Omega} \|\mathbb{P}(W_t = \cdot \mid W_0 = \underline{\sigma}) - \pi(\cdot)\|_{\text{TV}} < \epsilon\}, \quad t_{\text{mix}} \equiv t_{\text{mix}}(1/4),$$

where $\|\mu - \nu\|_{\text{TV}} \equiv \frac{1}{2}\sum_{\underline{\sigma} \in \Omega} |\mu(\underline{\sigma}) - \nu(\underline{\sigma})|$. In what follows it is convenient to consider the continuous-time Glauber dynamics X_t defined as follows. Place at each site $v \in V$ an i.i.d. rate-one Poisson clock; at each clock ring, we update the associated vertex in the same manner as in the discrete-time chain. The mixing time of the continuous chain X_t can be defined similarly and we denote it by $t_{\text{mix}}^{\text{ct}}$. It is well-known (cf. [6, Thm. 20.3]) that the two mixing times $t_{\text{mix}}, t_{\text{mix}}^{\text{ct}}$ satisfy the following relation.

Proposition 2.1. *Under the notation above,* $t_{\text{mix}}(\epsilon) \leqslant 4|V| t_{\text{mix}}^{\text{ct}}(\epsilon/2)$.

2.2. **Update sequence and grand coupling.** The *update sequence* along an interval $(t_0, t_1]$ is the set of tuples of the form (v, t, U), where $v \in V$ is the vertex to be updated, $t \in (t_0, t_1]$ is the update time, and $U \in \{0,1\}$ is the tentative update value of v ("tentative" as it might be an illegal update). An update (v, t, U) is said to be *blocked* (by hyperedge a) in configuration $\underline{\sigma}$, if $U = 1$ and there exists $a \in \partial v$ such that $\sigma(u) = 1$ for all $u \in \partial a \backslash \{v\}$.

Under this notation, the update rule of the Glauber dynamics can be rephrased as updating the spin at v to U at time t unless the update is blocked in X_t. Therefore $X_t^{\underline{\sigma}}$, the continuous-time chain starting from initial configuration $\underline{\sigma}$, can be expressed as a deterministic function of $\underline{\sigma}$ and the update sequence ξ. We will denote this function by $\mathbf{X}$ and write

$$X_t^{\underline{\sigma}} = \mathbf{X}[\xi, \underline{\sigma}; 0, t].$$

We remark that $\mathbf{X}$ depends on the underlying graph G implicitly.

A related update function $\mathbf{Y}$ is given by setting the spin at v to U at each update (v, t, U) *regardless of* whether the update is blocked or not. We define a family of processes $Y_{s,t}$ on the state space of 2^V which, given the all-one initial configuration $\underline{1}$ and the update sequence $\xi_{s,t}$ along the interval $(s, t]$, satisfies

$$Y_{s,t} \equiv \underline{1}, \text{ if } t \leqslant s, \quad \text{and } Y_{s,t} \equiv \mathbf{Y}[\xi_{s,t}, \underline{1}; s, t], \text{ if } t > s.$$

In the continuous-time setting, the update sequence $\xi \equiv ((v_\ell, t_\ell, U_\ell))_{\ell \geqslant 1}$ follows a marked poisson process where $\xi^\circ \equiv ((v_\ell, t_\ell))_{\ell \geqslant 1}$ is a poisson point process on $V \times [0, \infty)$ with rate 1 (per site) and $(U_\ell)_{\ell \geqslant 1}$ is a sequence of i.i.d. Bernoulli(1/2) random variables independent of $((v_\ell, t_\ell))_{\ell \geqslant 1}$. Using the same marked Poisson process ξ for different update functions, the discussion above provides a *grand coupling* of the processes $(X_t^{\underline{\sigma}})_{t \geqslant 0}$ and $(Y_{s,t})_{t \geqslant 0}$ for all possible values of $s, t, \underline{\sigma}$ simultaneously.

It is straightforward to check that, under the above setting, for any fixed $s \geqslant 0$ the process $(Z_t^{(s)})_{t \geqslant 0}$, where $Z_t^{(s)} \equiv Y_{s,s+t}$, is a continuous-time simple random walk on the hypercube $\{0,1\}^V$ with initial state $Y_{s,s} = \underline{1}$, in which each co-ordinate is updated at rate 1.

The purpose of introducing $Y_{s,t}$ is to utilize the monotonicity in the constraints of independent set and provide a uniform upper bound to $X_t^{\underline{\sigma}}$ for all $\underline{\sigma} \in \Omega$. For simplicity of notation, we write $Y_t \equiv Y_{0,t}$. For any pair of vectors X, Y in $\{0,1\}^V$, write $X \leqslant Y$ if and only if $X(v) \leqslant Y(v)$ for all $v \in V$.

Proposition 2.2. *Under the notations above, for all $\underline{\sigma} \in \Omega(G)$ and $s, t \geqslant 0$, we have*

$$X_t^{\underline{\sigma}} \leqslant Y_t \leqslant Y_{s,t}. \tag{1}$$

Proof. We proceed by showing that (1) holds for each configuration $\sigma \in \Omega(G)$ and update sequence $\xi = ((v_\ell, t_\ell, U_\ell))_{\ell \geqslant 1}$. By the right continuity of the process, it is enough to verify (1) at time 0 and the times of updates $(t_\ell)_{\ell \geqslant 1}$. When referring to the second inequality, we may assume in addition that $s \leqslant t$, as otherwise $Y_{s,t} = \underline{1}$ and thus there is nothing to prove.

At time $t_0 \equiv 0$, (1) holds since $\underline{\sigma} \leqslant \underline{1}$ for all $\underline{\sigma} \in \Omega(G)$. Suppose by induction we have verified (1) at all update times $(t_{\ell'})_{\ell' \leqslant \ell-1}$. Since there is no update between $t_{\ell-1}$ and t_ℓ, (1) remains true till the moment immediately before t_ℓ. At time $t = t_\ell$, the inequality is preserved if we successfully update v_ℓ to U_ℓ in each of the configurations $X_t^{\underline{\sigma}}, Y_t, Y_{s,t}$. If the update fails in some of the configurations, then (v_ℓ, t_ℓ, U_ℓ) must be blocked in $\lim_{\epsilon \downarrow 0} X_{t-\epsilon}^{\underline{\sigma}}$. In which case $U_\ell = 1$ and we set $X_t^{\underline{\sigma}}(v)$ to be 0, while $Y_t(v)$ and $Y_{s,t}(v)$ to be 1, again preserving the inequality. Combining the two cases together complete the induction hypothesis the ℓ'th update. □

Let $t_{\text{coup}} \equiv t_{\text{coup}}(\xi)$ be the time the grand coupling succeeds under updating sequence ξ:

$$t_{\text{coup}} \equiv \min\{t : \forall \underline{\sigma}, \underline{\tau} \in \Omega, X_t^{\underline{\sigma}} = X_t^{\underline{\tau}}\}.$$

A standard argument (cf. [6, Thm. 5.2]) implies that for all $t > 0$

$$\max_{\underline{\sigma} \in \Omega} \|P^t(\underline{\sigma}, \cdot) - \pi\|_{\text{TV}} \leqslant \mathbb{P}(t_{\text{coup}} > t), \tag{2}$$

where $P^t(\underline{\sigma}, \cdot)$ is the distribution at time t of the continuous-time chain, started from $\underline{\sigma}$.

2.3. **Discrepancy sequence and activation time.** In this section we take a closer look at the update process backward in time and in particular show how a discrepancy at time t can be traced back to discrepancies at earlier times. This will provide a necessary condition for $\{t_{\text{coup}} > t\}$.

Given an update sequence ξ, a vertex $v \in V$ and $t > 0$, let $t_-(v,t) \equiv t_-(v,t;\xi)$ be the time of the last update in ξ at v before time t. More explicitly, we define

$$t_-(v,t) \equiv 0 \vee \sup\{t' < t : (v,t',0) \in \xi \text{ or } (v,t',1) \in \xi\}.$$

Fix an update sequence ξ and time $t > 0$. If $t_{\text{coup}} = t_{\text{coup}}(\xi) > t$, then there must exist two initial configurations $\underline{\sigma}, \underline{\tau} \in \Omega$ and a "discrepancy" v_0 at time t, i.e., a vertex $v_0 \in V$ such that $X_t^{\sigma}(v_0) \neq X_t^{\tau}(v_0)$. Now we choose an arbitrary such discrepancy v_0, look at the last update of v_0 before time $t_1 \equiv t$ and denote its time by $t_0 \equiv t_-(v_0, t_1)$.

Assume without loss of generality that $X_{t_1}^{\sigma}(v_0) = 0 \neq 1 = X_{t_1}^{\tau}(v_0)$. To end up with a discrepancy at v_0 after the update at t_0, its tentative update value must be 1 and it must be blocked in $X_{t_0}^{\sigma}$ but not in $X_{t_0}^{\tau}$. Hence there must exists a hyperedge $a_0 \in \partial v_0$ such that

$$X_{t_0}^{\sigma}(\partial a_0 \backslash \{v_0\}) = \underline{1} \neq X_{t_0}^{\tau}(\partial a_0 \backslash \{v_0\}).$$

Consequently, there must exist at least one vertex $u \in \partial a_0 \backslash \{v_0\}$ at which the two configurations disagree at time t_0, namely,

$$X_{t_0}^{\sigma}(u) = 1 \neq 0 = X_{t_0}^{\tau}(u).$$

We arbitrarily choose one such vertex $u \in \partial a_0 \backslash \{v_0\}$ and denote it by v_{-1}.

Now apply the same reasoning for the update at v_{-1} at time $t_{-1} \equiv t_-(v_{-1}, t_0)$. We can find a hyperedge $a_{-1} \in \partial v_{-1}$ blocking the update $(v_{-1}, t_{-1}, 1) \in \xi$ in exactly one of the two configurations $X_{t_{-1}}^{\sigma}$ and $X_{t_{-1}}^{\tau}$ (namely, in the latter). Moreover, there must exists a discrepancy at a certain vertex $v_{-2} \in \partial a_{-1}$ at time t_{-1}. Repeating the process until time 0 produces a sequence of tuples $\zeta^\circ \equiv ((v_{-\ell}, t_{-\ell}, a_{-\ell}))_{0 \leqslant \ell \leqslant L}$, where $(t_\ell)_{-L \leqslant \ell \leqslant 0}$ satisfies that

$$t_1 = t, \quad t_{-L} = 0 \quad \text{and for all } 0 \leqslant \ell < L, \quad t_{-\ell} \equiv t_-(v_{-\ell}, t_{-(\ell-1)}) \leqslant t_{-(\ell-1)};$$

(v_{-L}, a_{-L}) satisfies that $a_{-L} = a_{-L+1}$ and

$$X_{t_{-L}}^{\sigma}(v_L) = X_0^{\sigma}(v_L) \neq X_0^{\tau}(v_L) = X_{t_{-L}}^{\tau}(v_L); \tag{3}$$

and for each $0 \leqslant \ell < L$:

1. The update $(v_{-\ell}, t_{-\ell}, 1)$ exists in ξ.
2. The hyperedge $a_{-\ell}$ contains $v_{-\ell}$ and $v_{-(\ell+1)}$ and the update $(v_{-\ell}, t_{-\ell}, 1)$ is blocked by $a_{-\ell}$ in exactly one of the two configurations $X_{t_{-\ell}}^{\sigma}$ and $X_{t_{-\ell}}^{\tau}$.
3. The vertex $v_{-\ell}$ is not updated in the time interval $(t_{-\ell}, t_{-(\ell-1)}]$.

Condition 2 in the above description is hard to analyse directly, because in general it is hard to control the probability that an update is blocked for the process X_t^{σ} at time t. This is where we use monotonicity ((1)). Observe that whenever an update $(v,t,1)$ is blocked by a hyper-edge a in one of the two processes X_t^{σ}, X_t^{τ} at time t, one of them must be all 1 on $\partial a \backslash v$, and by monotonicity so is $Y_t(\partial a \backslash v)$. Namely

$$\underline{1} \geqslant Y_t(\partial a \backslash v) \geqslant \max\{X_t^{\sigma}(\partial a \backslash v), X_t^{\tau}(\partial a \backslash v)\} = \underline{1}.$$

Meanwhile, since we are trying to update the value at v to 1 at time t, after this update (which is always successful in Y_t), $Y_t(v)$ equals to 1 as well. Therefore, the sequence ξ° satisfies that

$$Y_{t_{-\ell}}(\partial a_{-\ell}) = \underline{1}, \text{ for all } 0 \leqslant \ell < L. \tag{4}$$

Equation (4) is the key property of our proof and will be used repeatedly in what comes.

For convenience of later application, it is useful to consider also the following representation of ζ° with non-negative indices, which moves forward in time rather than backwards as in the original construction of ζ°: Let $\zeta \equiv ((v'_\ell, t'_\ell, a'_\ell))_{0 \leqslant \ell \leqslant L}$, be defined as

$$(v'_\ell, t'_\ell, a'_\ell) \equiv (v_{\ell-L}, t_{\ell-L}, a_{\ell-L}) \qquad \text{for each } 0 \leqslant \ell \leqslant L,$$

and write $t'_{L+1} \equiv t_1$ for the endpoint of the time interval. We will refer to such a sequence ζ as a *discrepancy sequence up to time* $t_1 = t$ (with respect to $\underline{\sigma}, \underline{\tau}$ and ξ). It is straightforward to check that

Lemma 2.3. *Given an update sequence ξ and a time $t \geqslant 0$, if $\{t_{\text{coup}} > t\}$, then there exists a discrepancy sequence ζ up to time t as defined above.*

We end the section with one more definition.

Definition 2.4. Let $t \geqslant 0$, $a \in F$ and $v \in \partial a$. We say that (v, a) is *activated* (resp. *s-activated*) at time t if the update sequence ξ contains an update $(v, t, 1)$ and $Y_t(\partial a) = \underline{1}$ (resp. $Y_{s,t}(\partial a) = \underline{1}$). We say that a got activated (resp. *s-activated*) at time t if (u, a) got activated (resp. *s-activated*) at time t for some $u \in \partial a$. We further define (v, a) to be *active* at time 0 for all $a \in F$, $v \in \partial a$.

3. A simplified proof of a weaker version

To illustrate the key ideas of our proof technique, we first define an auxiliary site percolation on the space-time slab of the update history, and use it to prove the following weaker version of Theorem 1.1.

Theorem 3.1. *For every n-vertex hypergraph G of edge size k and maximal degree Δ, the Glauber dynamics mixes in $O(n \log n)$ time if the graph satisfies one of the following conditions:*

1. *$\Delta \leqslant 2^{k/2}/(\sqrt{8}k^2)$.*
2. *$\Delta \leqslant 2^k/(9k^3)$ and G is linear.*

3.1. **The auxiliary percolation process.** We break up the space-time slab into time intervals $([T_i, T_{i+1}))_{i \geqslant 0}$ of length k, where $T_i \equiv ik$, $i = 0, 1, \ldots$ We shall neglect the possibility that a certain edge got activated precisely at some time T_i, as this has probability 0. For $t > 0$, define $i(t) \equiv \lfloor t/k \rfloor$. Let $\tilde{G}_F$ be an oriented graph with vertex-set $F \times \mathbb{N}$ and edge-set $\tilde{E}_F$ satisfying that for any pairs of hyperedges $a, b \in F$ and integers $i, j \in \mathbb{N}$, there is an (oriented) edge from site (a, i) to site (b, j) in $\tilde{E}_F$ if and only if

$$\partial a \cap \partial b \neq \varnothing \text{ and } j - i \in \{0, 1\}. \tag{5}$$

Definition 3.2. Fix an update sequence ξ. We say that a site (a, i) is *active* if $i = 0$ or a is T_{i-1}-activated at some time $t \in [T_i, T_{i+1})$. We say that a site (a, i) is *susceptible* if there exists $v \in \partial a$ such that v is not updated during the time interval $[T_i, T_{i+1})$. A site (a, i) is then called *bad* if either it is active or there exists $0 \leqslant j < i$ such that (a, j) is active and (a, ℓ) is susceptible for all $j + 1 \leqslant \ell \leqslant i$.

An example of the site percolation is given in Figure 1a where the underlying graph $\tilde{G}_F$ is given in Figure 1b. The set of bad sites can be viewed as a site percolation on the graph $\tilde{G}_F$ in which each site (a, i) is open if it is bad with respect to the update sequence ξ. The next lemma relates the success of the grand coupling to the existence of open paths in the aforementioned site percolation.

Lemma 3.3. *For every update sequence ξ satisfying $t_{\mathrm{coup}} > T_{M+1} = (M+1)k$, there exists an oriented path of sites in $\tilde{G}_F$ that starts from $F \times \{0\}$, ends at $F \times \{M\}$ and satisfies that every site along the path is bad with respect to ξ.*

Proof. Fix an update sequence ξ and an integer $M \in \mathbb{N}$ such that $t_{\mathrm{coup}} > T_{M+1}$. By the discussion in Section 2.3, we can find two initial configurations $\underline{\sigma}, \underline{\tau} \in \Omega(G)$ and a discrepancy sequence $\zeta^\circ \equiv ((v_{-\ell}, t_{-\ell}, a_{-\ell}))_{0 \leqslant \ell \leqslant L}$ from time $t_{-L} = 0$ up to time $t_1 = T_{M+1}$ with respect to $\underline{\sigma}, \underline{\tau}$ and ξ. We proceed to construct a path $\gamma \equiv \gamma(\zeta)$ in $\tilde{G}_F$ based on ζ°. (See Figure 1 for an illustration.)

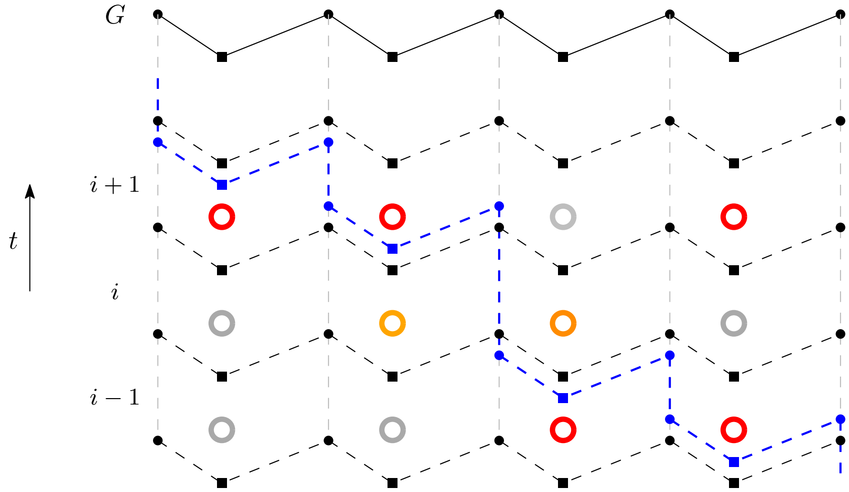
Recall that $i(t) = \lfloor t/k \rfloor$ is the time interval t belongs to. Naturally, we would like our path γ to pass through sites $\{(a_{-\ell}, i(t_{-\ell}))\}_{0 \leqslant \ell \leqslant L}$, where the updates in ζ° take place, and stays at each hyperedge until the next update happens at a nearby hyperedge. Let γ_ℓ denote the segment of γ corresponding to the ℓ'th update $(\nu_{-\ell}, a_{-\ell}, t_{-\ell})$ of ζ°. If the next (i.e., $(\ell-1)$'th) update happens at the same time interval as the ℓ'th update, i.e., $i(t_{-\ell}) = i(t_{-\ell+1})$, then we define γ_ℓ to be a singleton pair $((a_{-\ell}, i(t_{-\ell})))$. Otherwise if $i(t_{-\ell}) < i(t_{-\ell+1})$, we define γ_ℓ as the vertical line

$$\gamma_\ell \equiv \gamma_\ell(\zeta^\circ) \equiv ((a_{-\ell}, j) : i(t_{-\ell}) \leqslant j < i(t_{-\ell+1}))\,.$$

Let $\gamma \equiv \cup_{0 \leqslant \ell \leqslant L} \gamma_\ell$ be the sequential concatenation of $(\gamma_\ell)_{0 \leqslant \ell \leqslant L}$. It is easy to observe that each γ_ℓ is a connected path in $\tilde{G}_F$, γ_0 intersects $F \times \{M\}$ at $(a_0, i(t_1) - 1) = (a_0, M)$, and γ_L intersects $F \times \{0\}$ at $(a_{-L}, 0)$. To verify the rest of the requirements of Lemma 3.3, we first check that every site $(a, i) \in \gamma$ is bad, distinguishing three cases:

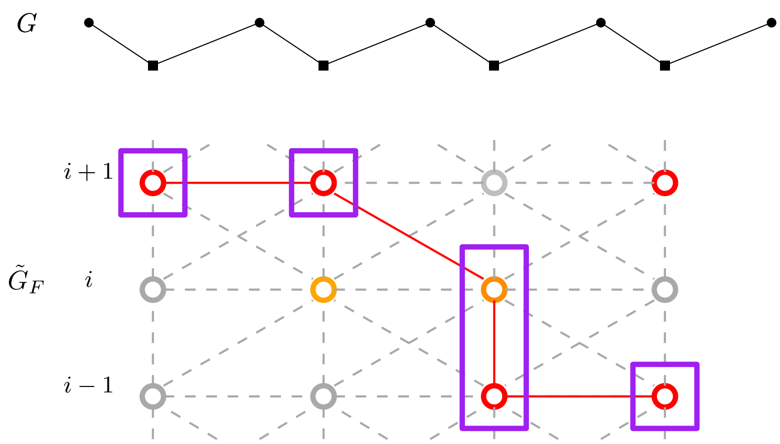
1. For $0 \leqslant \ell \leqslant L$, if γ_ℓ is a singleton, then it must have the form $\gamma_\ell = \{(a, i) = (a_{-\ell}, i(t_{-\ell}))\}$. By condition (4) $Y_{t_{-\ell}}(\partial a_{-\ell}) = \underline{1}$ and hence also $Y_{T_{i-1}, t_{-\ell}}(\partial a_{-\ell}) = \underline{1}$, implying that (a, i) is indeed active.
2. For $0 \leqslant \ell < L$, if γ_ℓ consists of more than one site, then arguing as above we know that the first site $(a_{-\ell}, i(t_{-\ell}))$ is active. For each of the remaining sites (a, i) with $i(t_{-\ell}) < i < i(t_{-(\ell-1)})$, rewriting the assumption gives
$$t_{-\ell} < T_i < T_{i+1} \leqslant t_{-(\ell-1)},$$
i.e., $v_{-\ell} \in \partial a_{-\ell}$ is not updated during the time interval $[T_i, T_{i+1})$ and so $(a_{-\ell}, i) \in \gamma_\ell$ is susceptible. Following the second case of Definition 3.2, $(a_{-\ell}, i)$ is bad for all $i(t_{-\ell}) < i < i(t_{-(\ell-1)})$.
3. For each $(a_{-L}, i) \in \gamma_L$, recall from the definition of ζ° that $a_{-L} = a_{-(L-1)}$ and that $t_-(v_{-L}, t_{-(L-1)}) = 0$, i.e., v_{-L} is never updated before $T_{i(t_{-(L-1)})}$. Hence all sites (a_{-L}, i) with $i < i(t_{-(L-1)})$ are susceptible. Their badness then follows from the fact that $(a_{-L}, 0)$ is bad (recall that $(a, 0)$ is defined to be active for all $a \in F$).

All that is left is to check that γ_ℓ is connected to $\gamma_{\ell-1}$ for each $1 \leqslant \ell \leqslant L$. The connectivity of γ_L to γ_{L-1} is trivial. For $1 \leqslant \ell < L$, the first site in $\gamma_{\ell-1}$ is $(a_{-(\ell-1)}, i_{\ell-1}^{\mathrm{start}} \equiv i(t_{-(\ell-1)}))$ and



(A) The discrepancy sequence projected onto $\tilde{G}_F$.

The dashed blue line represents the discrepancy sequence ζ, the red (resp. orange, gray) circles represent active (resp. susceptible, other) sites in $\tilde{G}_F$ and the red path represents γ. Note that the second site on the second row is susceptible but not bad since it is not immediately above any bad sites.



(B) An open path in the site percolation on $\tilde{G}_F$

The dashed gray lines are the underlining graph $\tilde{G}_F$ where we ignore the direction of the edges. The red line is the openpath γ constructed based on the discrepancy path ζ in panel A. And the purple rectangles marks the segments γ_ℓ.

FIGURE 1. Discrepancy sequence and site percolation

the last site in γ_ℓ is $(a_{-\ell}, i_\ell^{\text{end}})$ with

$$i_\ell^{\text{end}} \equiv \begin{cases} i(t_{-(\ell-1)}) - 1 & \text{if } i(t_{-(\ell-1)}) > i(t_{-\ell}), \\ i(t_{-\ell}) & \text{if } i(t_{-(\ell-1)}) = i(t_{-\ell}) \end{cases}.$$

In either case, site $(a_{-\ell}, i_\ell^{\text{end}})$ is connected to $(a_{-(\ell-1)}, i_{\ell-1}^{\text{start}})$ in $\tilde{G}_F$ since $\partial a_{-\ell} \cap \partial a_{-(\ell-1)} \supseteq \{v_{-\ell}\} \neq \varnothing$ and

$$i_{\ell-1}^{\text{start}} - i_\ell^{\text{end}} = 1 - \mathbf{1}\{i(t_{-(\ell-1)}) = i(t_{-\ell})\} \in \{0, 1\}.$$

$\square$

3.2. **Proof of Theorem 3.1.** By Proposition 2.1 and (2), we aim to show that under the assumptions of Theorems 3.1, there exists some constant C such that for $M \equiv \lceil C \log n \rceil$, $\mathbb{P}[t_{\text{coup}} > T_{M+1}] \leqslant 1/8$. In fact, as will be clear from the proof, once we find an M for which our analysis yields that $\mathbb{P}[t_{\text{coup}} > T_{M+1}] \leqslant 1/8$, doubling it gives $\mathbb{P}[t_{\text{coup}} > T_{2(M+1)}] \leqslant O(n^{-1})$.

Using Lemma 3.3, it suffices to bound the probability that there exists an oriented path from $F \times \{0\}$ consisting of bad sites. We begin with some basic properties of the auxiliary percolation process.

Proposition 3.4. *Let $\mathsf{A}(a,i)$ (resp. $\mathsf{S}(a,i)$) be the event that site (a,i) is active (resp. susceptible). Then for every distinct $(a,i),(b,j) \in \tilde{G}_F$ with $i \geqslant j \geqslant 1$,*

$$\mathbb{P}(\mathsf{A}(a,i)) \leqslant (k^2+1)2^{-k} + ke^{-k}, \quad \mathbb{P}(\mathsf{S}(a,i+1)) \leqslant ke^{-k} \tag{6}$$

$$\mathbb{P}(\mathsf{A}(a,i) \mid \mathsf{A}(b,j)) \leqslant (k^2+1)2^{-k+|\partial a \cap \partial b|} + ke^{-k}, \tag{7}$$

Moreover for every set of sites S in $\tilde{G}_F$, if for all $(b,j) \in S$, (a,i) and (b,j) are not connected in $\tilde{E}_F$, then $\mathsf{A}(a,i), \mathsf{S}(a,i)$ are independent of the events $\{\mathsf{A}(b,j), \mathsf{S}(b,j) : (b,j) \in S\}$.

Proof. The second part of (6) is simply a union bound. To show the first part of (6), we notice that if $\mathsf{A}(a,i)$ happens then so does one of the following scenarios

(a) Some $v \in \partial a$ is not updated in $[T_{i-1}, T_i)$, namely $\mathsf{S}(a, i-1)$ happens.
(b) Case (a) fails but a gets T_{i-1}-activated at some time $s \in [T_i, T_{i+1})$.

Case (a) happens with probability at most ke^{-k}. Conditioned on the failure of case (a), namely every $v \in \partial a$ being updated in $[T_{i-1}, T_i)$, $(Y_{T_{i-1},s}(u))_{u \in \partial a}$ become i.i.d. Bernoulli(1/2) r.v.'s for all $T_i \leqslant s < T_{i+1}$. For case (b) to happen, there must exist $u \in \partial a$ and $s \in [T_i, T_{i+1})$ such that $(u, s, 1) \in \xi$ and $Y_{T_{i-1},s}(\partial a) = \underline{1}$. Hence conditioned on the failure of case a, by Markov's inequality the probability of case (b) is at most $k^2 2^{-k}$, where the term k^2 represents the expected number of updates of the vertices in ∂a during $[T_i, T_{i+1})$ and 2^{-k} is the probability that $Y_{T_{i-1},s}(\partial a) = \underline{1}$ at the time of update. Combining the two cases gives the first half of (6). The proof of (7) is completely analogous, where the only difference is that we first argue that

$$\mathbb{P}(\mathsf{A}(a,i) \mid \mathsf{A}(b,j)) \leqslant \mathbb{P}(\mathsf{A}(a,i) \mid Y_{T_{i-1},s}(\partial a \cap \partial b) = \underline{1} \text{ for all } s \in [T_i, T_{i+1})).$$

and then apply the same reasoning as before to $\partial a \backslash \partial b$.

Finally, note that the events $\mathsf{A}(a,i)$ and $\mathsf{S}(a,i)$ depend on ξ only through $\partial a \times (T_{i-1}, T_{i+1})$. Hence the independency result follows from the independency of Poisson point process. $\square$

In order to perform a first moment calculation in an efficient manner we restrict our attention to a special type of path.

Definition 3.5. We say that an oriented path $((a_0,i_0),(a_1,i_1),\ldots,(a_r,i_r))$ in $\tilde{G}_F$ is a *minimal path* if $i_0 = 0$, $i_1 = 1$ and for all $j_1 \leqslant r-2$ and $j_2 \in [j_1+2, r]$ we have that $((a_{j_1}, i_{j_1}), (a_{j_2}, i_{j_2})) \notin \tilde{E}_F$. Let $\Gamma_{\min,r}$ be the collection of all minimal paths of length r.

Here we require $i_1 = 1$ because we are primarily interested in upperbounding the number of open paths connecting $F \times \{0\}$ to $F \times M$ for some $M \geqslant 1$ and additonal steps within $F \times \{0\}$ will not help. Observe that every oriented path in $\tilde{G}_F$ can be transformed into a minimal path by deleting some vertices from it.

Proof of Theorem 3.1, Part 1. Fix $M \equiv \lceil C \log n \rceil$ where the constant C shall be determined later. Note that by Lemma 3.3, if $t_{\mathrm{coup}} > T_{M+1}$, then there must be some minimal path

$$\gamma_M = ((a_0, i_0), (a_1, i_1), \ldots, (a_M, i_M)) \in \Gamma_{\min,M}$$

consisting of bad sites. We now estimate the expected number of such paths. For brevity, we call a minimal path γ bad if every site of γ is bad.

Using the notations of Proposition 3.4, for every $\gamma \in \Gamma_{\min,2r}$, we can write

$$\mathbb{E}[\mathbf{1}\{\gamma \text{ is bad}\}] \leqslant \mathbb{P}\Big(\bigcap_{\ell=0}^{2r} (\mathsf{A}(a_\ell, i_\ell) \cup \mathsf{S}(a_\ell, i_\ell))\Big) \leqslant \mathbb{P}\Big(\bigcap_{\ell=1}^{r} (\mathsf{A}(a_{2\ell}, i_{2\ell}) \cup \mathsf{S}(a_{2\ell}, i_{2\ell}))\Big),$$

where in the last step we discard all odd events in order to obtain the desirable independence. Indeed, by the definition of a minimal path, for all $j < \ell$, $(a_{2\ell}, i_{2\ell})$ is not connected to (a_{2j}, i_{2j}). Hence by Proposition 3.4, the events $\{\mathsf{A}(a_{2\ell}, i_{2\ell}) \cup \mathsf{S}(a_{2\ell}, i_{2\ell})\}_{1 \leqslant \ell \leqslant r}$ are mutually independent and

$$\mathbb{E}[\mathbf{1}\{\gamma \text{ is bad}\}] \leqslant \prod_{\ell=1}^{r} \left[(k^2+1)2^{-k} + 2ke^{-k}\right] \leqslant \left(2k^2 2^{-k}\right)^r \tag{8}$$

To conclude the proof we note that

$$|\Gamma_{\min,2r}| \leqslant n(2k(\Delta-1)+1)^{2r} \leqslant (4k^2\Delta^2)^r n, \tag{9}$$

where the term n accounts for the choice of the initial site of the path. By the above analysis the expected number of paths in $\Gamma_{\min,2r}$ consisting of bad sites is at most $(4k^2\Delta^2 \times 2k^2 2^{-k})^r n$. By our assumption that $\Delta < 2^{k/2}/(\sqrt{8}k^2)$, we get that there exists some constant C such that for $r = \lceil C \log n \rceil$ the last expectation is at most $1/8$. This concludes the proof of part 1 of Theorem 3.1. □

We now explain the necessary adaptations for the proof of part 2. In the new setup if $a \neq b$ and $\partial a \cap \partial b \neq \varnothing$, then $|\partial a \cap \partial b| = 1$. Thus the event that (a, i) is bad barely affects the probability that (b, j) is bad (for $a \neq b$). However, it is more challenging to control the conditional probability that $(a, i+1)$ is bad, given that (a, i) is bad. To overcome this difficulty we modify the underlying graph $\tilde{G}_F$ slightly.

Definition 3.6. Let $\bar{G}_F$ be an oriented graph with vertex-set $F \times \mathbb{N}$ and edge-set $\bar{E}_F$ satisfying that for every pairs of hyperedges $a, b \in F$ and integers $i, j \in \mathbb{N}$, there is an (oriented) edge from site (a, i) to site (b, j) in $\bar{E}_F$ if and only if $a = b$ and $j = i + 2$ or $a \neq b$ and

$$\partial a \cap \partial b \neq \varnothing \text{ and } j - i \in \{0, 1, 2\}. \tag{10}$$

We say that an oriented path $((a_0, i_0), (a_1, i_1), \ldots, (a_r, i_r))$ in $\bar{G}_F$ is a *minimal path* if $i_0 = 1$, $i_1 > 1$ and for all $j_1 \leqslant r - 2$ and $j_2 \in [j_1 + 2, r]$ we have that $((a_{j_1}, i_{j_1}), (a_{j_2}, i_{j_2})) \notin \bar{E}_F$. Let $\bar{\Gamma}_{\min,r}$ be the collection of all minimal paths of length r in $\bar{G}_F$.

Observe that every path γ in $\tilde{G}_F$ can be transformed into a path in $\bar{G}_F$ by deleting some of its vertices. Namely, whenever we have two consecutive steps in γ such that $((a_\ell, i_\ell), (a_{\ell+1}, i_{\ell+1})) \in \tilde{E}_F \backslash \bar{E}_F$, it must satisfy that $(a_{\ell+1}, i_{\ell+1}) = (a_\ell, i_\ell + 1)$ and one can check that in this case $((a_\ell, i_\ell), (a_{\ell+2}, i_{\ell+2})) \in \bar{G}_F$. By repeatedly deleting $(a_{\ell+1}, i_{\ell+1})$ from γ, where ℓ is the minimal index such that $(a_{\ell+1}, i_{\ell+1}) = (a_\ell, i_\ell + 1)$ and (a_ℓ, i_ℓ) has not been deleted already, we obtain a path in $\bar{G}_F$. From there, one can further take a subpath such that it is a minimal path in $\bar{G}_F$.

Proof of Theorem 3.1, Part 2. As before, if $t_{\text{coup}} > T_{2(M+1)}$, then there must be some minimal path $((a_0, i_0), (a_1, i_1), \ldots, (a_M, i_M))$ in $\bar{G}_F$ consisting of bad sites. We argue that the conditional probability that (a_ℓ, i_ℓ) is bad, given that $(a_0, i_0), (a_1, i_1), \ldots (a_{\ell-1}, i_{\ell-1})$ are all bad, is at most

$$(1 + k^2)2^{-(k-1)} + 2ke^{-k} \leqslant 3k^2 2^{-k} \quad (\text{for } k > 2).$$

Indeed, by the linearity assumption we have that either $(a_{\ell-1}, i_{\ell-1}) = (a_\ell, i_\ell - 2)$ or $|\partial a_{\ell-1} \cap \partial a_\ell| = 1$. In the first case, the same independency argument as before shows that the conditional probability that (a_ℓ, i_ℓ) is also bad is at most $(1 + k^2)2^{-k} + 2e^{-k}$. For the second case, (a_ℓ, i_ℓ) must be active and the desired bound follows from (7).

As before, we conclude by noting that $|\bar{\Gamma}_{\min,r}| \leqslant n(3k(\Delta - 1) + 1)^r \leqslant (3k\Delta)^r n$ and so the expected number of paths in $\bar{\Gamma}_{\min,r}$ consisting of bad sites is at most

$$(3k\Delta \times 3k^2 2^{-k})^r n = (9k^3 \Delta 2^{-k})^r n \leqslant 1/8,$$

provided that $r \geqslant C \log n$. □

4. General Hypergraphs

4.1. **Minimal path.** In this subsection we give the general setup for Theorem 1.1 and 1.3. The first improvement from Section 3 is based on the observation that although it takes time k to update all vertices of a hyperedge at least once with high probability, on average it only takes $O(1)$ time to update each vertex. This motivates us to study the propagation of discrepancies on both vertices and hyperedges in continuous time.

Recall the definition of activation from Definition 2.4. Observe that, in order for a discrepancy to propagate from an active hyperedge to a nearby hyperedge, the latter hyperedge must be activated before every vertex in their intersection is updated at least once, erasing all possible dependence. Thus we define the following continuous time analog of the time block from the previous section.

Definition 4.1. For each $v \in V$, $t \geqslant 0$, we define the *deactivation time* of v (after time t) as the first time v is updated after time t, namely,

$$T_+(v; t) \equiv \inf \{s > t : (v, s, 1) \in \xi \text{ or } (v, s, 0) \in \xi\},$$

For each $a, b \in F$ and $t \geqslant 0$, we define the *relative deactivation time* of a w.r.t. b (after time t) as the first time $s > t$ such that every vertex in the intersection $\partial a \cap \partial b$ is updated at least once by time s, namely,

$$T_+(a; b, t) \equiv \max_{v \in \partial a} T_+(v; b, t), \qquad \text{where } T_+(v; b, t) \equiv \begin{cases} T_+(v; t) & v \in \partial b, \\ t & v \notin \partial b \end{cases}.$$

In particular, the deactivation time of a is $T_+(a; t) \equiv T_+(a; a, t) = \max_{v \in \partial a} T_+(v; t)$.

Under the definition above, discrepancies can only pass from one hyperedge to another before the first hyperedge is relatively deactivated w.r.t. the latter hyperedge. In other word, (relative) deactivation time gives the time window of discrepancy propagation.

Let $\tilde{G}$ be the Cartesian product of G and the time interval $[0,\infty)$. Namely for two sites $(w_1,t_1),(w_2,t_2)\in\tilde{V}\equiv(V\cup F)\times[0,\infty)$, there exists an (oriented) edge connecting (w_1,t_1) to (w_2,t_2) in $\tilde{G}$ if and only if

$$w_1=w_2,\ t_1\leqslant t_2\quad\text{or}\quad w_1\in\partial w_2,\ t_1=t_2.$$

The set $\tilde{G}$ can be viewed as the continuous version of the space-time slab.

Definition 4.2. Given a sequence $\gamma_L\equiv\{(v_\ell,a_\ell,t_\ell)\}_{0\leqslant\ell\leqslant L}$, we say that γ_L is a path of length L in $\tilde{G}$ if $t_0=0$, $v_0\in\partial a_0, a_0=a_1$ and for each $1\leqslant\ell\leqslant L$ we have that $t_{\ell-1}<t_\ell$, $v_\ell\in\partial a_\ell$ and $\partial a_{\ell-1}\cap\partial a_\ell\neq\varnothing$. We further say that γ_L is a path up to time t to indicate that $t_L<t$. Let Γ_L denote the set of paths of length L and $\Gamma_L(t)$ denote the subset of paths up to time t.

In the previous section, we defined an auxiliary percolation on the discrete space-time slab such that every discrepancy sequence can be projected onto an open path in the percolation. Without a good analog of the percolation in continuous time, we look at the analog of "open paths" directly, i.e., paths of $\Gamma_L(t)$ that can be interpreted as the projections of discrepancy sequences.

Definition 4.3. Fix an update sequence ξ and a path $\gamma_L=\{(v_\ell,a_\ell,t_\ell)_{0\leqslant\ell\leqslant L}\}\in\Gamma_L$. For each $1\leqslant\ell\leqslant L$, we say that (v_ℓ,a_ℓ,t_ℓ), the ℓth step of γ_L, is a *minimal step* of γ_L if all of the following six events hold:

$$\begin{aligned}
&\mathsf{A}_\ell\equiv\{(v_\ell,t_\ell,1)\in\xi\}, &&\mathsf{B}_\ell\equiv\{t_\ell\leqslant T_+(a_\ell;a_{\ell-1},t_{\ell-1})\},\\
&\mathsf{C}_\ell\equiv\begin{cases}\{t_\ell>T_+(a_\ell;a_\ell,0)\}\cap(\cap_{r=1}^{\ell-2}\{t_\ell>T_+(a_\ell;a_r,t_r)\}) & \ell\geqslant 2\\ \varnothing^c & \ell=1\end{cases}, &&\\
&\mathsf{D}_\ell^1\equiv\{Y_{t_\ell}((\partial a_\ell\cap\partial a_{\ell-1}^c)\backslash\{v_\ell\})=\underline{1}\}, &&\mathsf{D}_\ell^2\equiv\{Y_{t_\ell}((\partial a_\ell\cap\partial a_{\ell-1})\backslash\{v_\ell\})=\underline{1}\},\\
&\mathsf{E}_\ell\equiv\begin{cases}\{Y_{t_\ell}(\partial a_{\ell-1}\backslash\partial a_\ell)\neq\underline{1}\} & v_\ell\in\partial a_{\ell-1}\cap\partial a_\ell, a_{\ell-1}\neq a_\ell\\ \varnothing^c & \text{otherwise}\end{cases}. &&
\end{aligned}\tag{11}$$

We say that γ_L is a *minimal path* if (v_ℓ,a_ℓ,t_ℓ) is a minimal step, for each $1\leqslant\ell\leqslant L$. We further say that γ_L is a *minimal path up to time* t, if it is a minimal path and $t_L<t\leqslant T_+(v_L;t_L)$. Let $\Gamma_{\min,L}$ be the set of minimal paths of length L. Similarly, we define $\Gamma_{\min,L}(t)\equiv\{\gamma_L\in\Gamma_{\min,L}:\gamma_L$ is a minimal path up to time $t\}$. We note that both $\Gamma_{\min,L}$ and $\Gamma_{\min,L}(t)$ depend on the update sequence ξ.

Remark 4.4. The six events above can be roughly explained as the following:

(1) A_ℓ: There is an update to 1 at vertex v_ℓ at time t_ℓ.
(2) B_ℓ: At time t_ℓ, a_ℓ has not been deactivated from the step *immedaitely before* it (i.e., the activation of $a_{\ell-1}$ at time $t_{\ell-1}$).
(3) C_ℓ: At time t_ℓ, a_ℓ has been deactivated from all steps *at least two steps ago* (i.e., the activations of a_r at time t_r for $r\leqslant\ell-2$).
(4) $\mathsf{D}_\ell^1\cap\mathsf{D}_\ell^2$: At time t_ℓ, the configuration on $\partial a_\ell\backslash\{v_\ell\}$ is all 1. Thus update at v_ℓ is prone to be blocked by a_ℓ. Here we differentiate D_ℓ^1 and D_ℓ^2 because the conditional distribution of Y_{t_ℓ} on $\partial a_\ell\cap a_{\ell-1}$ and $\partial a_\ell\cap a_{\ell-1}^c$ are very different and are easier to be analysed separately.
(5) E_ℓ: If consider the ℓ'th step being $(v_\ell,a_{\ell-1},t_\ell)$ instead of (v_ℓ,a_ℓ,t_ℓ), then the new tuple $(v_\ell,a_{\ell-1},t_\ell)$ does not satisfy the first five events because it violates D_ℓ^2. This requirement ensures that we stay at the same hyperedge whenever possible.

In the definition above, event $\mathsf{A}_\ell \cap \mathsf{D}_\ell^1 \cap \mathsf{D}_\ell^2$ guarrantees that (v_ℓ, a_ℓ) is activated at time t_ℓ. The event B_ℓ guarrantees that a_ℓ has not been deactivated w.r.t. $a_{\ell-1}$ after time $t_{\ell-1}$. Together, events $\{\mathsf{A}_\ell, \mathsf{B}_\ell, \mathsf{D}_\ell^1, \mathsf{D}_\ell^2\}_{1\leqslant \ell\leqslant L}$ imply that γ_L can potentially be the projection of some discrepancy sequence ζ (as will be proved in the lemma below).

Further, the events $\{\mathsf{C}_\ell\}_{1\leqslant\ell\leqslant L}$ imply that no subpath of γ_L satisfies the same condition while $\{\mathsf{E}_\ell\}_{1\leqslant\ell\leqslant L}$ further require paths to stay at the same hyperedge whenever possible (this requirement is imposed to obtain a better control on the number of minimal paths), justifying the name "minimal".

Lemma 4.5. *For each update sequence ξ, if $t_{\mathrm{coup}} > T$, then there exists a constant $L \geqslant 0$ and a minimal path $\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L} \in \Gamma_{\min,L}(T)$.*

Proof. Recall the construction of a discrepancy sequence $\zeta = \zeta(\xi) = ((u_i, b_i, s_i))_{0\leqslant i\leqslant M}$ (here we use the representation moving forward in time, with $(b_0, s_0) = (b_1, 0)$). It satisfies for all $1 \leqslant i \leqslant M$, that (a) $u_{i-1}, u_i \in \partial b_i$, (b) $(u_i, s_i, 1) \in \xi$, (c) $Y_{s_i}(b_i) = \underline{1}$ and (d)

$$s_{i-1} < s_i \leqslant T_+(u_{i-1}; s_{i-1}), \quad \text{hence in particular,} \quad s_{i-1} < s_i \leqslant T_+(b_i; b_{i-1}, s_{i-1}). \tag{12}$$

One can construct a minimal path based on ζ as follows:

1. Let $\gamma^1 \equiv ((v_\ell^1, a_\ell^1, t_\ell^1))_{0\leqslant\ell\leqslant 1} \equiv ((u_\ell, b_\ell, s_\ell))_{0\leqslant\ell\leqslant 1}$. It follows from the construction of a discrepancy sequence, in particular from the fact that $b_0 = b_1$, that γ^1 is a minimal path of length $L_1 \equiv 1$ up to time s_2.
2. For $2 \leqslant i \leqslant M$, suppose we have already constructed $\gamma^{i-1} = ((v_\ell^{i-1}, a_\ell^{i-1}, t_\ell^{i-1}))_{0\leqslant\ell\leqslant L_{i-1}} \in \Gamma_{\min,L_{i-1}}(s_i)$ with $L_{i-1} \leqslant i-1$ and
$$(v_{L_{i-1}}^{i-1}, t_{L_{i-1}}^{i-1}) = (u_{i-1}, t_{i-1}). \tag{13}$$
To construct γ^i, let
$$\ell_\star \equiv \ell_\star(i) \equiv \begin{cases} 0 & s_i \leqslant T_+(b_i; b_i, 0) \\ \min_{1\leqslant\ell\leqslant L_{i-1}}\{\ell : s_i \leqslant T_+(b_i; a_\ell^{i-1}, t_\ell^{i-1})\} & \text{otherwise} \end{cases}.$$
Note that $\ell_\star$ is well-defined since by (12) and (13), the condition $\{\ell : s_i \leqslant T_+(b_i; a_\ell^{i-1}, t_\ell^{i-1})\}$ is satisfied by $\ell = L_{i-1}$. If $\ell_\star = 0$, then there exists $v_\star \in \partial b_i$ such that $s_i \leqslant T_+(v_\star; 0)$. In this case, we define $L_i = 1$ and
$$\gamma^i \equiv ((v_\ell^i, a_\ell^i, t_\ell^i))_{0\leqslant\ell\leqslant 1} \equiv ((v_\star, b_i, 0), (u_i, b_i, s_i)).$$
Otherwise, we define $L_i = \ell_\star + 1$, $(v_\ell^i, a_\ell^i, t_\ell^i) \equiv (v_\ell^{i-1}, a_\ell^{i-1}, t_\ell^{i-1}), 0 \leqslant \ell \leqslant \ell_\star$ and
$$(v_{L_i}^i, a_{L_i}^i, t_{L_i}^i) \equiv \begin{cases} (u_i, a_{\ell_\star}^{i-1}, s_i) & \text{if } u_i \in \partial a_{\ell_\star}^{i-1} \text{ and } Y_{s_i}(\partial a_{\ell_\star}^{i-1}) = \underline{1}, \\ (u_i, b_i, s_i) & \text{otherwise.} \end{cases} \tag{14}$$
In either case, one can check that the six events defined in (11) are satisfied for $\ell = L_i$ and hence $(v_{L_i}^i, a_{L_i}^i, t_{L_i}^i)$ is a minimal step of γ^i. Indeed, the occurrence of $\mathsf{A}_{L_i}, \mathsf{B}_{L_i}, \mathsf{D}_{L_i}^1, \mathsf{D}_{L_i}^2$ follows from the construction of a discrepancy sequence, the occurrence of C_{L_i} follows from the minimality of $\ell_\star$ and that of E_{L_i} follows from (14). By construction, $((v_r^i, a_r^i, t_r^i))_{0\leqslant\ell\leqslant L_i-1}$ is a subpath of $\gamma^{i-1} \in \Gamma_{\min,L_{i-1}}(s_i)$ and hence is minimal path itself. Therefore γ^i is a minimal path of length L_i up to s_{i+1}.

To conclude the proof, one can take γ^M and note that $T_+(v_{L_M}^M; t_{L_M}^M) = T_+(u_M; s_M) \geqslant T$ by the definition of ζ. ☐

Remark 4.6. We will use $\Gamma_{\text{proj},L} \subset \Gamma_{\min,L}$ to denote the set of minimal paths that are projected from some discrepancy sequences as in the proof above.

Lemma 4.5 implies that for every time $T \geqslant 0$ and integer $L \geqslant 1$,

$$\mathbb{P}(t_{\text{coup}} \geqslant T) \leqslant \mathbb{P}(\Gamma_{\text{proj},L-1}(T) \neq \varnothing) + \mathbb{P}(\Gamma_{\text{proj},L} \neq \varnothing) \tag{15}$$

$$\leqslant \mathbb{P}(\Gamma_{\text{proj},L-1}(T) \neq \varnothing) + \mathbb{P}(\Gamma_{\min,L} \neq \varnothing). \tag{16}$$

The next lemma bounds the first term on the right hand side.

Lemma 4.7. *Let $T \in \mathbb{N}$. Denote $L = \lfloor cT \rfloor$, where $c = \frac{1}{4\log(2k^2\Delta)}$. Then*

$$\mathbb{P}(\Gamma_{\text{proj},L-1}(T) \neq \varnothing) \leqslant n\exp(-T/4).$$

Proof. Consider constructing a path in $\Gamma_{\text{proj},L-1}(T)$ by first choosing the locations of "jumps" and then picking their times. The number of ways of choosing a sequence $\underline{w} \equiv (a_\ell)_{0\leqslant \ell\leqslant L-1}$ such that $\partial a_\ell \cap \partial a_{\ell-1} \neq \varnothing$ for all $1 \leqslant \ell \leqslant L-1$, is at most $n(\Delta k)^L$. For each fixed $\underline{w}$, recursively define $T_0 \equiv 0$ and $T_{\ell+1} \equiv T_+(a_{\ell+1}; a_\ell, T_\ell)$ for $0 \leqslant \ell \leqslant L-2$. If there exists a sequences of times $\underline{s} \equiv (s_\ell)_{0\leqslant\ell\leqslant L-1}$ and vertices $\underline{v} \equiv (v_\ell)_{0\leqslant\ell\leqslant L-1}$ such that the path $\gamma_{L-1} \equiv ((v_\ell, a_\ell, s_\ell))_{0\leqslant\ell\leqslant L-1}$ is a minimal path up to time T, then $s_0 = T_0 = 0$, and one can show inductively that

$$s_\ell \leqslant T_+(a_\ell; a_{\ell-1}, s_{\ell-1}) \leqslant T_+(a_\ell; a_{\ell-1}, T_{\ell-1}) = T_\ell,$$

using the induction step ($s_{\ell-1} \leqslant T_{\ell-1}$) and the monotonicity of $T_+(a; b, t)$ in t. The existence of $\underline{s}$ further implies that

$$T_L \equiv T_+(a_{L-1}; a_{L-1}, T_{L-1}) \geqslant T_+(v_{L-1}; s_{L-1}) \geqslant T.$$

By construction, the joint law of $(T_\ell - T_{\ell-1})_{0\leqslant\ell\leqslant L}$ is stochastically dominated by that of $(Z_\ell)_{0\leqslant\ell\leqslant L}$ where $Z_\ell := \max(Z_{\ell,1}, \dots, Z_{\ell,k})$ and $(Z_{\ell,i})_{1\leqslant i\leqslant k, 1\leqslant\ell\leqslant L}$ are i.i.d. exponential random variables with rate 1. Note that using the order statistic of $Z_{\ell,1}, \dots, Z_{\ell,k}$, we can decompose Z_ℓ into a sum $\sum_{i=1}^k J_{\ell,i}$ of independent exponential r.v.'s with $\mathbb{E}[J_{\ell,i}] = i^{-1}$. Hence for all $\lambda \in (0, 1/2)$ and $\ell \leqslant L$,

$$\mathbb{E}[e^{\lambda Z_\ell}] = \prod_{i=1}^k \mathbb{E}[e^{\lambda J_{\ell,i}}] = \prod_{i=1}^k \left(1 + \frac{\lambda}{i-\lambda}\right) \leqslant \prod_{i=1}^k \left(1 + \frac{1}{i}\right) \leqslant 2e^{\frac{1}{2}+\cdots+\frac{1}{k}} \leqslant e^{\log(2k)}.$$

By Markov's inequality, independence of Z_ℓ's and the aforementioned stochastic domination,

$$\mathbb{P}(T_L \geqslant T) \leqslant \mathbb{E}[e^{(T_L-T)/2}] = e^{-T/2}e^{L\log(2k)} \leqslant e^{-[1-2c\log(2k)]T/2}.$$

Therefore by the arguments above together with the choice $c = \frac{1}{4\log(2k^2\Delta)}$,

$$\mathbb{P}(\Gamma_{\text{proj},L-1}(T) \neq \varnothing) \leqslant n(\Delta k)^{cT}\mathbb{P}(T_L \geqslant T) \leqslant n(\Delta k)^{cT}e^{-[1-2c\log(2k)]T/2} = ne^{-T/4},$$

as desired. $\square$

4.2. **Redacted path.** In the remainder of the section, we bound the size of $\Gamma_{\min,L}$. Our basic approach is to bound for each $\gamma_L \in \Gamma_{\min,L}$ the expected number of ways to extend γ_L by two steps. However, the "vanilla-version" of this expectation can be much bigger than 1. Intuitively, for general hypergraphs, hyperedges may be highly overlapping with each other. Thus when one hyperedge is all 1 in process Y_t, its neighbouring hyperedges may also be all 1 with little extra cost. This phenomenon leads to numerous local "tangles" where

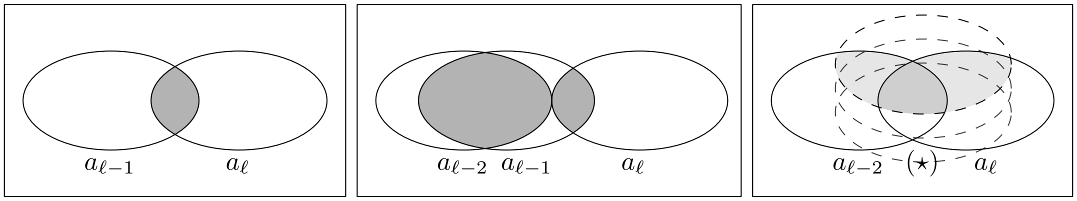


FIGURE 2. Good/Type-I/Type-II branching in redacted paths

multiple choices of the immediate next step exist for the same second-next step, blowing up the number of extensions significantly.

To overcome the aforementioned obstacle, we divide the steps of a minimal path into good branchings (i.e., small overlaps) and bad branchings (i.e., large overlaps) and skip the "tangles" by ignoring the first step of a bad branching and recording only the "key" step instead. More precisely, given a path $\gamma_L = \{(u_\ell, b_\ell, s_\ell)\}_{0\leqslant \ell\leqslant L} \in \Gamma_L$, we classify each of the steps $1 \leqslant \ell \leqslant L$ into one of the following three cases: (see also Figure 2)

(1) We say that (v_ℓ, a_ℓ, t_ℓ) is a good branching if $|\partial a_{\ell-1} \cap \partial a_\ell| \leqslant k/3$.

(2) We say that $\{(v_{\ell-1}, a_{\ell-1}, t_{\ell-1}), (v_\ell, a_\ell, t_\ell)\}$ is a type-I (bad) branching if

$$|\partial a_{\ell-2} \cap \partial a_{\ell-1}| > k/3, \quad \partial a_{\ell-2} \cap \partial a_\ell = \varnothing.$$

(3) We say that $\{(v_{\ell-1}, a_{\ell-1}, t_{\ell-1}), (v_\ell, a_\ell, t_\ell)\}$ is a type-II (bad) branching if

$$|\partial a_{\ell-2} \cap \partial a_{\ell-1}| > k/3, \quad \partial a_{\ell-2} \cap \partial a_\ell \neq \varnothing.$$

For each $\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant \ell\leqslant L} \in \Gamma_L$, we use the following greedy algorithm to partition γ_L into disjoint segments of size one or two such that all but (possibly) the last block satisfy one of the above three types of branchings (we assume $L > 1$ to avoid triviality).

1. Observe that (by Definition 4.2) $a_1 = a_0$ implies that $\partial a_2 \cap \partial a_0 \neq \varnothing$. Therefore $\{(v_1, a_1, t_1), (v_2, a_2, t_2)\}$ forms a type-II branching and is taken as a block in the partition.
2. For each $\ell \leqslant L-2$, if after a certain number of iterations we have partitioned the first ℓ steps of γ_L into one of the above three cases, then we look at the $(\ell+1)$'th step $(v_{\ell+1}, a_{\ell+1}, t_{\ell+1})$:
   - If $(v_{\ell+1}, a_{\ell+1}, t_{\ell+1})$ forms a good branching then we take $\{(v_{\ell+1}, a_{\ell+1}, t_{\ell+1})\}$ as a block of size 1 in the partition.
   - Otherwise, we take $\{(v_{\ell+1}, a_{\ell+1}, t_{\ell+1}), (v_{\ell+2}, a_{\ell+2}, t_{\ell+2})\}$ as a segment of size 2 in the partition. By definition, it must either form a type-I branching or a type-II branching.
3. If step 2 ends with a complete partition of the path (i.e., the last two steps form a type-I or type-II branching), then we are done. Otherwise, it must give a partition of the first $L-1$ steps of γ_L. To obtain a partition of the first L steps we then let (v_L, a_L, t_L) form a block of size 1, and call it a *half-branching* if it is not a good-branching itself.

Given the unique partition defined above, we define $I_{\mathrm{g}}, I_1, I_\star, I_2 \subseteq \{1, \ldots, L\} \equiv [L]$, each as a function of $\gamma_L \in \Gamma_L$, as the set of (the indices of) steps that belong to good branchings, type-I branchings, the first step of type-II branchings and the second step of type-II branchings, respectively, in the partition of γ_L by the above procedure. The four sets $I_{\mathrm{g}}, I_1, I_\star, I_2$ form a partition of $[L]$ or $[L-1]$. In particular, $1 \in I_\star$, $2 \in I_2$ and L belongs to none of the four sets if the last step is a half-branching.

Definition 4.8. Given a sequence $\tilde{\gamma}_L \equiv \{(v_\ell, a_\ell, t_\ell)\}_{0\leqslant \ell\leqslant L}$, we say that $\tilde{\gamma}_L$ is a *redacted path* in $\tilde{G}$ of length L if there exists path $\gamma_L = \{(u_\ell, b_\ell, s_\ell)\}_{0\leqslant\ell\leqslant L} \in \Gamma_L$ such that for each $0 \leqslant \ell \leqslant L$,

$$(v_\ell, a_\ell, t_\ell) = \begin{cases} (u_\ell, b_\ell, t_\ell), & \ell \in \{0, L\} \cup I_{\mathrm{g}}(\gamma_L) \cup I_1(\gamma_L) \cup I_2(\gamma_L) \\ (\star, \star, \star), & \ell \in I_\star(\gamma_L) \end{cases}. \tag{17}$$

We will refer to steps equaling to $(\star, \star, \star)$ as *redacted steps*.

With some abuse of notation, we define $I_{\mathrm{g}} \equiv I_{\mathrm{g}}(\tilde{\gamma}_L) \equiv I_{\mathrm{g}}(\gamma_L)$ and define $I_1, I_2, I_\star$ similarly. Here we remark that for redacted path $\tilde{\gamma}_L$, the set $I_\star$ (resp. I_2) can be defined as the collection of redacted steps (resp. the collection of steps following a redacted step) and thus does not depend on the choice of γ_L. Let $M(\tilde{\gamma}_L) \equiv |I_{\mathrm{g}}| + \frac{1}{2}|I_1| + |I_2|$ be the *branching length* of $\tilde{\gamma}_L$, where the factor $\frac{1}{2}$ is taken because we want each type-I branching to contribute $+1$ to $M(\tilde{\gamma}_L)$. We define

$$\begin{aligned} \tilde{\Gamma}_M &\equiv \{\tilde{\gamma}_L \in \tilde{\Gamma} : M(\tilde{\gamma}_L) = M, \quad L \in I_{\mathrm{g}} \cup I_1 \cup I_2\}, \\ \tilde{\Gamma}_{M+1/2} &\equiv \{\tilde{\gamma}_L \in \tilde{\Gamma} : M(\tilde{\gamma}_L) = M, \quad L \notin I_{\mathrm{g}} \cup I_1 \cup I_2\}, \end{aligned}$$

to be the set of redacted paths of branching length M that end with regular branchings or half-branchings respectively. We finally denote the set of all redacted paths by $\tilde{\Gamma}$.

In the definition above, redacted steps, i.e., steps equaling to $(\star, \star, \star)$, can never appear twice in a row, and the last step of $\tilde{\gamma}_L$ is never redacted. Let

$$I_{\mathrm{m}} \equiv I_{\mathrm{m}}(\tilde{\gamma}_L) \equiv \{\ell : \ell, \ell - 1 \notin I_\star(\tilde{\gamma}_L)\} = \begin{cases} I_{\mathrm{g}} \cup I_1 & L \in I_2 \\ I_{\mathrm{g}} \cup I_1 \cup \{L\} & \text{otherwise} \end{cases}.$$

Using the convention that $T_+(v; \star, \star) = T_+(a; \star, \star) = 0$ for all $v \in V$ and $a \in F$, we can extend the definition of a minimal step to redacted paths.

Definition 4.9. We say that (v_ℓ, a_ℓ, t_ℓ) is a *minimal step* in $\tilde{\gamma}_L = ((v_i, a_i, t_i))_{0\leqslant i\leqslant L}$ if and only if $\ell \in I_{\mathrm{m}}$ and (v_ℓ, a_ℓ, t_ℓ) satisfies the six events of (11). For $\ell \in I_2$ we say that (v_ℓ, a_ℓ, t_ℓ) is a *minimal type-II step*, if all of the following events hold:

$$\begin{aligned} &\mathsf{A}_\ell \equiv \{(v_\ell, t_\ell, 1) \in \xi\}, \quad \widetilde{\mathsf{B}}_\ell \equiv \{t_\ell \leqslant T_+(a_\ell; T_+(a_{\ell-2}; t_{\ell-2}))\}, \\ &\mathsf{C}_\ell \equiv \{t_\ell > T_+(a_\ell; a_\ell, 0)\} \cap (\cap_{r=1}^{\ell-2}\{t_\ell > T_+(a_\ell; a_r, t_r)\}), \\ &\widetilde{\mathsf{D}}_\ell \equiv \{Y_{t_\ell}(\partial a_\ell \backslash \{v_\ell\}) = \underline{1}\}. \end{aligned} \tag{18}$$

Definition 4.10. Fix an update sequence ξ and a redacted path $\tilde{\gamma}_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L} \in \tilde{\Gamma}$. We say that $\tilde{\gamma}_L \in \tilde{\Gamma}$ is a *minimal redacted path* if,

(1) For each $\ell \in I_{\mathrm{m}}$, (v_ℓ, a_ℓ, t_ℓ) is a minimal step of $\tilde{\gamma}_L$.
(2) For each $\ell \in I_2$, (v_ℓ, a_ℓ, t_ℓ) is a minimal type-II step of $\tilde{\gamma}_L$.

We denote the set of minimal redacted paths by $\tilde{\Gamma}_{\min}$ and define $\tilde{\Gamma}_{\min,M} \equiv \{\tilde{\gamma} \in \tilde{\Gamma}_{\min} : \tilde{\gamma}$ has branching length $M\}$ where M takes value from $\frac{1}{2}\mathbb{N}$.

Lemma 4.11. *For every update sequence ξ and length $L \geqslant 0$, if $\Gamma_{\min,L} \neq \varnothing$, then $\tilde{\Gamma}_{\min,M} \cup \tilde{\Gamma}_{\min,M+1/2} \neq \varnothing$ where $M = \lfloor (L-1)/2 \rfloor$.*

Proof. For each $\gamma_L = ((u_\ell, b_\ell, s_\ell))_{0\leqslant\ell\leqslant L} \in \Gamma_{\min,L}$, we can construct $\tilde{\gamma}_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L}$ according to (17). It is straight forward to check that $\tilde{\gamma}_L$ is a redacted path with branching length $M(\tilde{\gamma}_L) \geqslant (L-1)/2$ and for all $\ell \in I_{\mathrm{m}}$ we have that (v_ℓ, a_ℓ, t_ℓ) is a minimal step of $\tilde{\gamma}_L$. We now verify that for all $\ell \in I_2$ we have that (v_ℓ, a_ℓ, t_ℓ) is a minimal type-II step of $\tilde{\gamma}_L$. To do so, it suffices to verify that if $\ell \in I_2$ then $\widetilde{\mathsf{B}}_\ell$ holds (the occurrence of $\mathsf{A}_\ell, \mathsf{B}_\ell$ and $\widetilde{\mathsf{D}}_\ell$ follows from the fact that $\gamma_L \in \Gamma_{\min,L}$). Fix some $\gamma_L \in \Gamma_{\min,L}$ and $\ell \in I_2$. From (11) we know that events $\mathsf{B}_{\ell-1}$ and B_ℓ hold. It follows that

$$\begin{aligned} t_{\ell-1} &\leqslant T_+(a_{\ell-1}; a_{\ell-2}, t_{\ell-2}) \leqslant T_+(a_{\ell-2}; t_{\ell-2}), \\ t_\ell &\leqslant T_+(a_\ell; a_{\ell-1}, t_{\ell-1}) \quad \leqslant T_+(a_\ell; t_{\ell-1}). \end{aligned}$$

Combining the last two equations and using the monotonicity of $T_+(a;t)$ in t yields that

$$t_\ell \leqslant T_+(a_\ell; T_+(a_{\ell-2}; t_{\ell-2})).$$

Consequently, $\tilde\gamma_L \in \tilde\Gamma_{\min}$. Truncating $\tilde\gamma_L$ such that $M = \lfloor (L-1)/2 \rfloor$ concludes the proof. □

4.3. **Recursion of two steps.** In this subsection we complete the recursion on $\tilde\Gamma_{\min,M}$ and conclude the proof of Theorem 1.1. The main idea is to bound the expected number of ways of extending a redacted path by each one of the three types of branchings. Since every step of a good branching or a type-I branching is also a minimal step, we split the discussion according to minimal steps and minimal type-II steps. For each $a \in F$, let $\mathbb{N}(a) \equiv \{b \in F : \partial a \cap \partial b \neq \varnothing\}$ be the hyperedge-neighbourhood of a. Define

$$\mathbb{N}^{>}(a) \equiv \{b \in \mathbb{N}(a) : |\partial a \cap \partial b| > k/3\}, \quad \mathbb{N}^{\leqslant}(a) \equiv \mathbb{N}(a) \backslash \mathbb{N}^{>}(a).$$

Throughout the section, we will use $\tilde\gamma_L$ to denote the redacted paths of length L (but varying branching length).

Fix a path $\tilde\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L} \in \tilde\Gamma$ and a vertex-hyperedge pair (v,a) satisfying $a \in \mathbb{N}(a_L)$, $v \in \partial a$. We denote by $\tilde\gamma^{\mathrm{m}}_{L+1}(v,a,t) = ((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L+1}$ the redacted path extended from $\tilde\gamma_L$ with $(v_{L+1}, a_{L+1}, t_{L+1}) \equiv (v,a,t)$. We define

$$N^{\mathrm{m}}_L \equiv N^{\mathrm{m}}(\tilde\gamma_L; v, a) \equiv |\{t : \tilde\gamma^{\mathrm{m}}_{L+1}(v,a,t) \in \tilde\Gamma_{\min},\ L+1 \in I_{\mathrm{m}}(\tilde\gamma^{\mathrm{m}}_{L+1}(v,a,t))\}| \tag{19}$$

to count the number of possible minimal steps using (v,a). Note that N^{m}_L is a.s. finite since it is bounded from above by the number of updates at v between time t_L and $T_+(a_{L+1}; a_L, t_L)$, which in turn is a.s. finite. We further use $\tilde\gamma^{\mathrm{II}}_{L+2}(v,a,t) = ((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L+2}$ to denote the redacted path extended from $\tilde\gamma_L$ with

$$(v_{L+1}, a_{L+1}, t_{L+1}) \equiv (\star, \star, \star), \quad (v_{L+2}, a_{L+2}, t_{L+2}) \equiv (v,a,t)$$

and define

$$N^{\mathrm{II}}_L \equiv N^{\mathrm{II}}(\tilde\gamma_L; v, a) \equiv |\{t : \gamma^{\mathrm{II}}_{L+2}(v,a,t) \in \tilde\Gamma_{\min,M+1},\ L+2 \in I_2(\tilde\gamma^{\mathrm{II}}_{L+2}(v,a,t))\}|.$$

Finally, for each integer $M \geqslant M(\tilde\gamma_L)$, we let $\tilde\Gamma_{\min,M}(\tilde\gamma_L)$ be the collection of redacted paths in $\tilde\Gamma_{\min,M}$ that agree with $\tilde\gamma_L$ in their first L steps and write $N_{\min,M}(\tilde\gamma_L) := |\tilde\Gamma_{\min,M}(\tilde\gamma_L)|$.

Lemma 4.12. *For every integer $L \geqslant 1$, $\tilde\gamma_L \in \tilde\Gamma$, $a \in \mathbb{N}(a_L)$ and $v \in \partial a$,*

$$\mathbb{E}[N^{\mathrm{m}}(\tilde\gamma_L; v, a) \mid \tilde\gamma_L \in \tilde\Gamma_{\min}] \leqslant C_{\mathrm{m}} \begin{cases} \frac{1}{k} & a = a_L, \\ \frac{1}{m_{L+1}+1} 2^{-(k-1-m_{L+1})} & a \neq a_L, v \notin a_L, \\ \frac{k-m_{L+1}}{k} \frac{1}{m_{L+1}+1} 2^{-(k-1-m_{L+1})} & a \neq a_L, v \in a_L, \end{cases}$$

where $m_{L+1} \equiv |\partial a \cap \partial a_L \backslash \{v\}|$ and C_{m} is an absolute constant independent of Δ, k.

Lemma 4.13. *For every integers $L \geqslant M \geqslant 1$, $\tilde\gamma_L \in \tilde\Gamma_M$, $a \in \mathbb{N}(a_L)$ and $v \in \partial a$,*

$$\mathbb{E}[N^{\mathrm{II}}(\tilde\gamma_L; v, a) \mid \tilde\gamma_L \in \tilde\Gamma_{\min,M}] \leqslant 2^{-(k-1)}(1 + \log k).$$

The proof of Lemma 4.12 and Lemma 4.13 is postponed to Section 4.4 and Section 4.5, respectively. We first apply both lemmas to derive the main result

Theorem 4.14. *For all integers $L \geqslant M \geqslant 1$, and redacted path $\tilde{\gamma}_L \in \tilde{\Gamma}_M$,*

$$\mathbb{E}[N_{\min,M+1}(\tilde{\gamma}_L) \mid \tilde{\gamma}_L \in \tilde{\Gamma}_{\min,M}] \leqslant C \begin{cases} \Delta k^2 2^{-k} & G \text{ is linear} \\ \Delta^2 2^{-k} + \Delta k^2 2^{-2k/3} & \text{otherwise} \end{cases},$$

where C is an absolute constant independent of Δ, k.

Proof. Fix some $\tilde{\gamma}_L = ((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L} \in \tilde{\Gamma}_M$. For brevity define $\tilde{\mathbb{E}}_L \equiv \mathbb{E}[\cdot \mid \tilde{\gamma}_L \in \tilde{\Gamma}_{\min,M}] = \mathbb{E}[\cdot \mid \tilde{\gamma}_L \in \tilde{\Gamma}_{\min}]$. Let $N^{\mathrm{g}}_{\min,M+1}(\tilde{\gamma}_L)$ be the number of redacted paths in $\tilde{\Gamma}_{\min,M+1}(\gamma_L)$ such that their last block is a good branching and define $N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}_L)$ and $N^{\mathrm{II}}_{\min,M+1}(\tilde{\gamma}_L)$ similarly. By the construction of redacted paths,

$$N_{\min,M+1}(\tilde{\gamma}_L) = N^{\mathrm{g}}_{\min,M+1}(\tilde{\gamma}_L) + N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}_L) + N^{\mathrm{II}}_{\min,M+1}(\tilde{\gamma}_L).$$

We bound the three cases separately:

1. We first bound $\tilde{\mathbb{E}}_L[N^{\mathrm{g}}_{\min,M+1}(\tilde{\gamma}_L)]$. For any two hyperedges $a, b \in F$, let $m(a,b) \equiv |\partial a \cap \partial b|$ be the size of their overlap. For each $a \in N^{\leqslant}(a_L)$, applying Lemma 4.12 to $v \in \partial a \backslash \partial a_L$ and $v \in \partial a \cap \partial a_L$ separately yields that
$$\sum_{v \in \partial a \backslash \partial a_L} \tilde{\mathbb{E}}_L N^{\mathrm{m}}(\tilde{\gamma}_L; v, a) \leqslant C_{\mathrm{m}} \cdot \frac{|\partial a \backslash \partial a_L|}{m(a, a_L) + 1} 2^{-(k-1-m(a,a_L))},$$
$$\sum_{v \in \partial a \cap \partial a_L} \tilde{\mathbb{E}}_L N^{\mathrm{m}}(\tilde{\gamma}_L; v, a) \leqslant C_{\mathrm{m}} \cdot |\partial a \cap \partial a_L| \cdot \frac{k+1-m(a,a_L)}{k m(a,a_L)} 2^{-(k-m(a,a_L))}.$$
Combining the two estimates yields the following upper bound on the number of good branchings,
$$\begin{aligned} \tilde{\mathbb{E}}_L[N^{\mathrm{g}}_{\min,M+1}(\tilde{\gamma}_L)] &= \sum_{a \in \mathbb{N}^{\leqslant}(a_L)} \sum_{v \in \partial a} \tilde{\mathbb{E}}_L[N^{\mathrm{m}}(\tilde{\gamma}_L; v, a)] \\ &\leqslant 3C_{\mathrm{m}} \sum_{a \in \mathbb{N}^{\leqslant}(a_L)} \frac{k+1-m(a,a_L)}{m(a,a_L)} 2^{-(k-m(a,a_L))} \leqslant 3C_{\mathrm{m}}(\Delta k) \cdot k 2^{-2k/3}. \end{aligned} \tag{20}$$
2. We now bound the expected number of type-I branchings. If $\tilde{\gamma}_{L+2} = ((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L+2} \in \tilde{\Gamma}_{\min,M+1}$ is extended from $\tilde{\gamma}_L$ via a type-I branching, then $\tilde{\gamma}_{L+1}$, defined as the first $(L+1)$ steps of $\tilde{\gamma}_{L+2}$, must have branching length $M + \frac{1}{2}$. We first enumerate the possible extensions of $\tilde{\gamma}_L$ (via a type-I branching) through some fixed redacted path $\tilde{\gamma}_{L+1} \equiv \tilde{\gamma}^{\mathrm{m}}_{L+1}(v_{L+1}, a_{L+1}, t_{L+1})$. Let
$$A_{L+1} \equiv A_{L+1}(\tilde{\gamma}_{L+1}) \equiv \mathbb{N}(a_{L+1}) \backslash \mathbb{N}(a_L)$$
be the set of a_{L+2}'s that together with a_{L+1} may form a type-I branching extended from $\tilde{\gamma}_L$. For each $a \in A_{L+1}$, we split the discussion into two cases, $v \in \partial a \backslash \partial a_{L+1}$ and $v \in \partial a \cap \partial a_{L+1}$. Define $\tilde{\mathbb{E}}_{L+1} \equiv \mathbb{E}[\cdot \mid \tilde{\gamma}_{L+1} \in \tilde{\Gamma}_{\min,M+1/2}]$. Applying similar reasonings as in the derivation

of (20), we get that

$$\tilde{\mathbb{E}}_{L+1}[N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}_{L+1})] \leqslant 3C_{\mathrm{m}} \sum_{a\in A_{L+1}} \frac{k+1-m(a,a_{L+1})}{m(a,a_{L+1})} 2^{-(k-m(a,a_{L+1}))} \equiv 3C_{\mathrm{m}} \sum_{a\in A_{L+1}} m(a,a_{L+1}) f(m(a,a_{L+1})), \tag{21}$$

where

$$f(m) \equiv \frac{k+1-m}{m^2} 2^{-(k-m)}.$$

The function $f(m)$ is an increasing function on $3 \leqslant m \leqslant k-1$ and $f(1), f(2) \leqslant 3f(3)$. Recall that $\partial a_{L+2} \cap \partial a_L = \varnothing$. Therefore the sizes of overlaps $\{m(a, a_L)\}_{a\in A_{L+1}}$ satisfy that

$$m(a,a_{L+1}) \leqslant |\partial a_{L+1} \backslash \partial a_L|, \quad \sum_{a\in A_{L+1}} m(a,a_{L+1}) \leqslant \Delta |\partial a_{L+1}\backslash \partial a_L| = \Delta(k - m(a_L, a_{L+1})).$$

Therefore,

$$\begin{aligned}
\tilde{\mathbb{E}}_{L+1}[N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}_{L+1})] &\leqslant 9C_{\mathrm{m}} f(k - m(a_L,a_{L+1})) \sum_{a\in A_{L+1}} m(a,a_{L+1}) \\
&= 9C_{\mathrm{m}} \Delta \frac{m(a_L,a_{L+1})+1}{k-m(a_L,a_{L+1})} 2^{-m(a_L,a_{L+1})}.
\end{aligned}$$

Now we sum over all possible choices of (v_{L+1}, a_{L+1}). Observe that if $a_{L+1} = a_L$, then $\partial a_L \cap \partial a_{L+2} \neq \varnothing$, leading to a type-II branching. Thus $A_L \equiv \mathbb{N}^{>}(a_L)\backslash\{a_L\}$ is the set of possible a_{L+1}'s. Observe from Lemma 4.12 that the upper bound of $N^{\mathrm{m}}(\tilde{\gamma}_L; v, a)$ does not depend on t_{L+1}. A similar calculation to (20) gives that

$$\begin{aligned}
\tilde{\mathbb{E}}_L[N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}_L)] &= \sum_{a\in A_L} \sum_{v\in\partial a} \tilde{\mathbb{E}}_L\big[N^{\mathrm{m}}(\tilde{\gamma}_L; v, a) \tilde{\mathbb{E}}_{L+1}[N^{\mathrm{I}}_{\min,M+1}(\tilde{\gamma}^{\mathrm{m}}_{L+1}(v,a,t))]\big] \\
&\leqslant 3C_{\mathrm{m}} \sum_{a\in A_L} \frac{k+1-m(a_L,a)}{m(a_L,a)} 2^{-(k-m(a_L,a))} \cdot 9C_{\mathrm{m}}\Delta \frac{m(a_L,a)+1}{k-m(a_L,a)} 2^{-m(a_L,a)} \\
&\leqslant 108 C_{\mathrm{m}}^2 \Delta 2^{-k} |A_L| \leqslant 324 C_{\mathrm{m}}^2 \Delta^2 k 2^{-k},
\end{aligned}$$

where the last step uses the fact that

$$|A_L| \leqslant \frac{\Delta k}{\min_{a\in\mathbb{N}^{>}(a_L)} m(a,a_L)} \leqslant 3\Delta.$$

3. Finally, using Lemma 4.13, we can bound the number of minimal type-II branchings:

$$\tilde{\mathbb{E}}_L N^{\mathrm{II}}_{\min,M+1} \leqslant \sum_{a\in\mathbb{N}(a_L)} \sum_{v\in\partial a} \tilde{\mathbb{E}}_L N^{\mathrm{II}}(\tilde{\gamma}_L; v, a) \leqslant (\Delta k^2) \cdot 2^{-(k-1)}(1+\log k).$$

Combining the three cases together completes the proof for general hypergraphs.

We now turn to linear hypergraphs. In this setup all branchings must be good (because $m(a,b) \leqslant 1$ for all $a,b \in F$). Therefore applying (20), we have

$$\begin{aligned}\tilde{\mathbb{E}}_L[N_{\min,M+1}(\tilde{\gamma}_L)] = \tilde{\mathbb{E}}_L[N^{\mathrm{g}}_{\min,M+1}(\tilde{\gamma}_L)] &= \sum_{a\in\mathbb{N}(a_L)}\sum_{v\in\partial a}\tilde{\mathbb{E}}_L[N^{\mathrm{m}}(\tilde{\gamma}_L;v,a)] \\ &\leqslant 3C_{\mathrm{m}}\sum_{a\in\mathbb{N}(a_L)}\frac{k+1-m(a,a_L)}{m(a,a_L)}2^{-(k-m(a,a_L))} = 6C_{\mathrm{m}}(\Delta k)\cdot k2^{-k}.\end{aligned}$$

This concludes the proof for linear hypergraphs. □

Proof of Theorem 1.1. The assertion of Theorem 1.1 follows by combining Lemma 4.5, (16), Lemma 4.7, Lemma 4.11 and Theorem 4.14. □

4.4. **Number of minimal steps.** In this subsection we prove Lemma 4.12. Throughout the section, we assume $\tilde{\gamma}_L \in \tilde{\Gamma}, a \in \mathbb{N}(a_L), v \in \partial a$ and for brevity of notation write $\tilde{\gamma}_{L+1}(t) \equiv \tilde{\gamma}^{\mathrm{m}}_{L+1}(v,a,t)$ (recall that $\tilde{\gamma}^{\mathrm{m}}_{L+1}(v,a,t) = ((v_\ell,a_\ell,t_\ell))_{0\leqslant\ell\leqslant L+1}$ is the redacted path extended from $\tilde{\gamma}_L$ with $(v_{L+1},a_{L+1},t_{L+1}) = (v,a,t)$). Recall the definitions of minimal branching and minimal type-II branching in (11) and (18). We define

$$\begin{aligned}&\mathsf{M}^{\mathrm{m}}_\ell(t) \equiv \mathsf{M}^{\mathrm{m}}_\ell(\tilde{\gamma}_{L+1}(t)) \equiv \mathsf{A}_\ell(t)\cap\mathsf{B}_\ell(t)\cap\mathsf{C}_\ell(t)\cap\mathsf{D}^1_\ell(t)\cap\mathsf{D}^2_\ell(t)\cap\mathsf{E}_\ell(t) \quad \forall\ell\in I_{\mathrm{m}}(\tilde{\gamma}_{L+1}(t)),\\ &\mathsf{M}^{\mathrm{II}}_\ell(t) \equiv \mathsf{M}^{\mathrm{II}}_\ell(\tilde{\gamma}_{L+1}(t)) \equiv \mathsf{A}_\ell(t)\cap\widetilde{\mathsf{B}}_\ell(t)\cap\mathsf{C}_\ell(t)\cap\widetilde{\mathsf{D}}_\ell(t) \quad \forall\ell\in I_2(\tilde{\gamma}_{L+1}(t)).\end{aligned}$$

The argument t, whose role is to indicate that $(v_{L+1},a_{L+1},t_{L+1}) = (v,a,t)$, is included as we shall soon vary t. However, we henceforth omit t from the notation for all events with $\ell \leqslant L$, as they do not depend on the value of t. We further write

$$\mathsf{N}_L \equiv \{\tilde{\gamma}_L \in \tilde{\Gamma}\} = \Big[\bigcap_{\ell\in I_{\mathrm{m}}}\mathsf{M}^{\mathrm{m}}_\ell\Big]\cap\Big[\bigcap_{\ell\in I_2}\mathsf{M}^{\mathrm{II}}_\ell\Big].$$

By Campbell's theorem,

$$\mathbb{E}[N^{\mathrm{m}}_L \mid \tilde{\gamma}_L \in \tilde{\Gamma}_{\min}] = \mathbb{E}\Big[\sum_{t:(v,t,1)\in\xi}\mathbf{1}\{\mathsf{M}^{\mathrm{m}}_{L+1}(t)\}\ \Big|\ \mathsf{N}_L\Big] = \frac{1}{2}\int_{t_L}^{\infty}\mathbb{P}(\mathsf{M}^{\mathrm{m}}_{L+1}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t))dt,$$

where $N^{\mathrm{m}}_L = N^{\mathrm{m}}(\tilde{\gamma}_L;v,a)$ is defined in (19). This motivates the following lemma.

Lemma 4.15. *Under the above notation,*

$$\mathbb{P}(\mathsf{M}^{\mathrm{m}}_{L+1}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t)) \leqslant 2^{-|\partial a\cap\partial a_L^c\setminus\{v\}|}\cdot\mathbb{P}(\mathsf{B}_{L+1}(t),\mathsf{D}^2_{L+1}(t),\mathsf{E}_{L+1}(t)\mid Y_{t_L}(\partial a_L)=\underline{1}).$$

Roughly speaking, the event $\mathsf{M}^{\mathrm{m}}_{L+1}(t)$ is contained in the intersection of two groups of events that are roughly independent: (1). the events $\mathsf{B}_{L+1}(t)$, $\mathsf{D}^2_{L+1}(t)$ and $\mathsf{E}_{L+1}(t)$ depending on ∂a_L. (2). the events $\mathsf{D}^1_{L+1}(t)$ and

$$\mathsf{C}'_{L+1}(t) \equiv \bigcap_{u\in\partial a\setminus\partial a_L}\Big[\{T_+(u;0)\leqslant t\}\cap\big(\cap_{\ell=1}^{L-1}\{T_+(u;a_\ell,t_\ell)\leqslant t\}\big)\Big] \supseteq \mathsf{C}_{L+1}(t). \tag{22}$$

depending on $\partial a\setminus\partial a_L$. The first group depends on N_L only through $Y_{t_{L-1}}(\partial a_L) = \underline{1}$, whereas for the second group, conditioning on $\mathsf{C}'_{L+1}(t)$, namely that every $u \in \partial a\setminus\partial a_L$ has been updated at least once since time 0 or its last apperance in $\tilde{\gamma}_L$, we intuitively expect that $\mathsf{D}^1_{L+1}(t) = \{Y_t(\partial a\cap\partial a_L^c\setminus\{v\}) = \underline{1}\}$ is roughly independent of everything else and happens with probability at most $2^{-|\partial a\cap\partial a_L^c\setminus\{v\}|}$.

In light of the above discussion, we expect that (omitting t's from the notation)

$$\begin{aligned}\mathbb{P}(\mathsf{M}^{\mathrm{m}}_{L+1} \mid \mathsf{N}_L, \mathsf{A}_{L+1}) &\lesssim \mathbb{P}(\mathsf{D}^1_{L+1} \mid \mathsf{C}'_{L+1}) \cdot \mathbb{P}(\mathsf{B}_{L+1}, \mathsf{D}^2_{L+1}, \mathsf{E}_{L+1} \mid Y_{t_L}(\partial a_L) = \underline{1}) \\ &\approx 2^{-|\partial a \cap \partial a_L^c \backslash \{v\}|} \cdot \mathbb{P}(\mathsf{B}_{L+1}, \mathsf{D}^2_{L+1}, \mathsf{E}_{L+1} \mid Y_{t_L}(\partial a_L) = \underline{1}).\end{aligned}$$

However, the event $\mathsf{N}_L = \{\tilde{\gamma}_L \in \tilde{\Gamma}_L\}$ may depend, through events E_ℓ and $\tilde{\mathsf{B}}_\ell$, on updates at a vertex after it last appears in some hyperedges of $\tilde{\gamma}_L$. While intuitively "additional updates" and also conditioning that the restriction of the configuration to certain hyperedges at certain times will not be all 1, "can only help", overcoming such dependencies is the main technical obstacle in the proof below. Through a subtle conditioning argument we will establish a positive correlation between the relevant events. For the sake of continuity of the argument, we postponed the proof of Lemma 4.15 to Section 4.6.

The last lemma we need before proving Lemma 4.12 concerns random walks on hypercubes. Its proof is also postponed to Section 4.6. Let $(Z_i)_{i\in\mathbb{Z}_+}$ be the (discrete-time) lazy simple random walk on the m-dimensional hypercube $\{0,1\}^m$ where in each step, a coordinate is chosen uniformly at random and updated to 0 or 1 with equal probability. Let $H_1 \equiv \inf\{i > 0 : Z_i = \underline{1}\}$ be the hitting time of $\underline{1}$ and let T_+ be the first time by which each coordinate which equals 1 at time 0 was updated at least once.

Lemma 4.16. *For every $m \geqslant 2$, the expected number of visits to $\underline{1}$ before T_+ satisfies*

$$\mathbb{E}\Big[\sum_{0\leqslant i<T_+} \mathbf{1}\{Z_i = \underline{1}\}\Big] \leqslant \begin{cases} \frac{6}{m} & Z_0 \neq (1,1,1,\dots,1) \\ 2+\frac{6}{m} & Z_0 = (1,1,1,\dots,1)\end{cases}. \tag{23}$$

We now prove Lemma 4.12.

Proof of Lemma 4.12. Recall that $m_{L+1} \equiv |\partial a \cap \partial a_L \backslash \{v\}|$. By Lemma 4.15 and the Campbell's theorem,

$$\begin{aligned}&2\mathbb{E}[N^{\mathrm{m}}(\tilde{\gamma}_L; v, a) \mid \tilde{\gamma}_L \in \tilde{\Gamma}_{\min}] \\ &\leqslant \int_{t_L}^{\infty} 2^{-|\partial a \cap \partial a_L^c \backslash \{v\}|}\mathbb{P}(\mathsf{B}_{L+1}(t), \mathsf{D}^2_{L+1}(t) \cap \mathsf{E}_{L+1}(t) \mid Y_{t_L}(\partial a_L) = \underline{1}) dt \\ &= 2^{-(k-1-m_{L+1})}\mathbb{E}\Big[\int_{t_L}^{T_+(a;a_L,t_L)} \mathbf{1}\{Y_t(\partial a_L) \in \Theta_{L+1}\} dt \,\Big|\, Y_{t_L}(\partial a_L) = \underline{1}\Big],\end{aligned} \tag{24}$$

where

$$\Theta_{L+1} \equiv \Theta_{L+1}(\tilde{\gamma}_L; v, a) \equiv \begin{cases} \{\sigma \in \{0,1\}^{\partial a_L} : \sigma_{\partial a_L \cap \partial a} = \underline{1}\}, & v \notin \partial a_L \\ \{\sigma \in \{0,1\}^{\partial a_L} : \sigma_{\partial a_L \cap \partial a \backslash \{v\}} = \underline{1}, \sigma_{\partial a_L \cap \partial a^c} \neq \underline{1}\}, & v \in \partial a_L\end{cases}$$

is the range of $Y_t(\partial a_L)$ restricted on $\mathsf{D}^2_{L+1}(t) \cap \mathsf{E}_{L+1}(t)$. We now apply the result of Lemma 4.16, differentiating the three cases:

1. $v \notin \partial a_L$: Let $(\tilde{Y}_i)_{i\in\mathbb{Z}_+}$ be the skeleton chain (i.e., the chain that records the configuration of Y_t on $\partial a_L \cap \partial a$ after every time a vertex in $\partial a_L \cap \partial a$ is updated) of the continuous time Markov chain $Y_t(\partial a_L \cap \partial a)$ starting from time t_L. Note that $(\tilde{Y}_i)_{i\in\mathbb{Z}_+}$ is a lazy simple random walk on the m_{L+1}-dimensional hypercube and that the time between two steps of $\tilde{Y}_i$ in $Y_t(\partial a_L \cap \partial a)$ are i.i.d. random variables with $\mathrm{Exp}(m_{L+1})$ distribution.

Therefore, we can rewrite the expectation on the RHS of (24) in terms of $\tilde{Y}_i$, namely,

$$\text{RHS of (24)} = 2^{-(k-1-m_{L+1})}\frac{1}{m_{L+1}}\mathbb{E}\Big[\sum_{0\leqslant i\leqslant \tilde{T}} \mathbf{1}\{\tilde{Y}_i = \underline{1}\} \;\Big|\; Y_0 = \underline{1}\Big],$$

where $\tilde{T}$ is the number of steps in $(\tilde{Y}_i)_{i\in\mathbb{Z}_+}$ until every vertex of $\partial a_L \cap \partial a$ is updated at least once. Applying Lemma 4.16, we get that

$$\text{RHS of (24)} \leqslant 2^{-(k-1-m_{L+1})}\frac{6}{m_{L+1}}.$$

2. $v \in \partial a_L, a \neq a_L$: Let $(\tilde{Z}_i)_{i\in Z_+}$ be the skeleton chain of $Y_t(\partial a_L)$ starting from t_L with $\tilde{Z}_0 = Y_{t_L}(\partial a_L) = \underline{1}$ and $T_0 \equiv \min\{i \geqslant 1 : \tilde{Z}_i \neq \underline{1}\}$ be the time of the first 0-update in $(\tilde{Z}_i)_{i\in Z_+}$. By the construction of Θ_{L+1}, for all $i \geqslant 1$, if $\tilde{Z}_i \in \Theta_{L+1}$ then we must have that $i \geqslant T_0$. For brevity of notation, let $A \equiv \partial a_L \cap \partial a\backslash\{v\}$ and define $\tilde{T}$ to be the number of steps in $(\tilde{Z}_i)_{i\in\mathbb{Z}_+}$ until which every vertex in $\partial a_L \cap \partial a = A \cup \{v\}$ is updated at least once. Observe that $\tilde{T}$ corresponds to the deactivation time $T_+(a; a_L, t_L)$ in the original process. By the strong Markov property and the total probability formula,

$$\begin{aligned}\text{RHS of (24)} &\leqslant 2^{-(k-1-m_{L+1})}\frac{1}{k}\mathbb{E}\Big[\sum_{T_0\leqslant i\leqslant \tilde{T}} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\} \;\Big|\; \tilde{Z}_0 = \underline{1}\Big] \\ &\leqslant 2^{-(k-1-m_{L+1})}\frac{1}{k}\mathbb{E}\Big[\mathbb{E}\Big[\sum_{T_0\leqslant i\leqslant \tilde{T}} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\} \;\Big|\; \tilde{Z}_{T_0}(A)\Big] \;\Big|\; \tilde{Z}_0 = \underline{1}\Big].\end{aligned}$$

If $m_{L+1} = 0$, meaning that v is the only vertex in the intersection of a_L and a, then $\tilde{T}$ simply follows the exponential distribution of rate 1 and

$$\text{RHS of (24)} \leqslant 1 \cdot 2^{-(k-1)}.$$

For $m_{L+1} \geqslant 1$, one can bound $\tilde{T}$ from above by $T_A + T_v$, where T_A is the number of steps until every vertex in A is updated at least once and

$$T_v \equiv \min\{i > T_A : v \text{ is updated at step } i\} - T_A$$

is the number of additional steps until v is updated for the first time after time T_A. It follows that

$$\sum_{T_0\leqslant i\leqslant \tilde{T}} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\} \leqslant \sum_{T_0\leqslant i\leqslant T_A} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\} + \sum_{T_A< i\leqslant T_A+T_v} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\}. \tag{25}$$

For the second summation of (25), observe that T_v follows the Geometric$(1/k)$ distribution and for any value of $\tilde{Z}_{T_0}(A)$ and $i \geqslant T_A$, we have that $\tilde{Z}_i(A)$ is uniformly distributed on $\{0,1\}^A$ with $|A| = m_{L+1}$. Therefore for all $\underline{z} \in \{0,1\}^A$

$$\mathbb{E}\Big[\sum_{T_A< i\leqslant T_A+T_v} \mathbf{1}\{\tilde{Z}_i(A) = \underline{1}\} \;\Big|\; \tilde{Z}_{T_0}(A) = \underline{z}\Big] \leqslant 2^{-m_{L+1}}\mathbb{E}T_v = k2^{-m_{L+1}}.$$

For the first sum on the RHS of (25), we split the discussion according to whether $\tilde{Z}_{T_0}(A) = \underline{1}$ or not. By the symmetry of the k vertices of ∂a_L, we get that

$$\mathbb{P}(\tilde{Z}_{T_0}(A) = \underline{1}) = (k - m_{L+1})/k.$$

Applying Lemma 4.16 to the restriction of $\tilde{Z}_i$ to A yields that

$$\mathbb{E}\Big[\sum_{T_0<i\leqslant T_A}\mathbf{1}\{\tilde{Z}_i(A)=\underline{1}\}\,\Big|\,\tilde{Z}_0=\underline{1}\Big]\leqslant\Big[\frac{k-m_{L+1}}{k}\cdot\Big(2+\frac{6}{m_{L+1}}\Big)+\frac{m_{L+1}}{k}\cdot\frac{6}{m_{L+1}}\Big]\frac{k}{m_{L+1}},$$

where the term k/m_{L+1} is obtained via Wald's equation, by noting that the time between two updates in A has a Geometric(m_{L+1}/k) distribution. Combining all pieces together, we have that for all $1\leqslant m_{L+1}\leqslant k-1$,

$$\begin{aligned}\text{RHS of (24)}&\leqslant 2^{-(k-1-m_{L+1})}\Big(\frac{1}{m_{L+1}}\frac{2(k-m_{L+1})+6}{k}+2^{-m_{L+1}}\Big)\\&\leqslant 2^{-(k-1-m_{L+1})}\frac{k-m_{L+1}}{k}\frac{9}{m_{L+1}}.\end{aligned}$$

3. $a_L=a$: This is similar to the second case and for $k\geqslant 3$

$$\text{RHS of (24)}\leqslant\frac{1}{k}\mathbb{E}\Big[\sum_{0\leqslant i\leqslant\tilde{T}}\mathbf{1}\{\tilde{Z}_i(A)=\underline{1}\}\,\Big|\,\tilde{Z}_0=\underline{1}\Big]\leqslant\frac{1}{k}\Big[2+\frac{6}{k}+k2^{-(k-1)}\Big]\leqslant\frac{5}{k},$$

where the two terms in the third step are obtained from an argument similar to (25).

Those three cases conclude the proof with $C_{\mathrm{m}}\equiv 9$. □

4.5. **Number of minimal type-II steps.** In this section we prove Lemma 4.13. Fix $\tilde{\gamma}_L\in\tilde{\Gamma}_M, a\in\mathbb{N}(a_L), v\in\partial a$ and write $\tilde{\gamma}_{L+2}(t)\equiv\tilde{\gamma}^{\mathrm{II}}_{L+2}(v,a,t)$. Recall the definition of $\mathsf{M}^{\mathrm{m}}_\ell$, $\mathsf{M}^{\mathrm{II}}_\ell$ and N_L from Section 4.4 and define $\mathsf{M}^{\mathrm{II}}_{L+2}(t)$ similarly. By Campbell's theorem,

$$\mathbb{E}[N^{\mathrm{II}}_L\mid\tilde{\gamma}_L\in\tilde{\Gamma}_{\min,M}]=\mathbb{E}\Big[\sum_{t:(v,t,1)\in\xi}\mathbf{1}\{\mathsf{M}^{\mathrm{II}}_{L+2}(t)\}\,\Big|\,\mathsf{N}_L\Big]=\frac{1}{2}\int_{t_L}^{\infty}\mathbb{P}(\mathsf{M}^{\mathrm{II}}_{L+2}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t))dt.$$

The next lemma is the type-II analog of Lemma 4.15, the proof of which is postponed to Section 4.6, after the introduction of relevant notations in the proof of Lemma 4.15.

Lemma 4.17. *Under the notations above, for all $\tilde{\gamma}_L\in\tilde{\Gamma}_M, v\in V, a\in F$ and $t>t_L$,*

$$\mathbb{P}(\mathsf{M}^{\mathrm{II}}_{L+2}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t))\leqslant 2^{-(k-1)}\mathbb{P}(\tilde{\mathsf{B}}_{L+2}(t)).$$

Proof of Lemma 4.13. By Lemma 4.13 and Campbell's theorem,

$$\begin{aligned}\mathbb{E}[N^{\mathrm{II}}_L\mid\tilde{\gamma}_L\in\tilde{\Gamma}_{\min}]&\leqslant 2^{-k}\int_{t_L}^{\infty}\mathbb{P}(\tilde{\mathsf{B}}_{L+2}(t))dt=2^{-k}\mathbb{E}[T_+(a;T_+(a_L;t_L))-t_L]\\&=2^{-k}\cdot 2\mathbb{E}[T_+(a_L;t_L)-t_L]=2^{-(k-1)}\sum_{1=1}^{k}\frac{1}{i}\leqslant 2^{-(k-1)}(1+\log k),\end{aligned}$$

where we have used the fact that the coupon collector time of k coupons is $\sum_{1=1}^{k}\frac{1}{i}$. □

4.6. **Remaining Lemmas.** In this subsection we complete the proof of Lemmas 4.15, 4.16 and 4.17.

Proof of Lemma 4.15. For each $u\in\partial a\cup\partial a_L$, let

$$\ell_-(u)\equiv 0\vee\max\{1\leqslant\ell\leqslant L:\ell\notin I_\star, u\in\partial a_\ell\}$$

be the last step in $\tilde{\gamma}_{L+1}$ such that the corresponding hyperedge contains u and let $t_-(u)\equiv t_{\ell_-(u)}$ be the time of that step. In particular, if u has never appeared in the previous steps, then $t_-(u) = 0$. We consider the set $S_{L+1} \equiv \cup_{u \in \partial a \cup \partial a_L} \{u\} \times (t_-(u), \infty)$, and (recalling $Y_0(V) = \underline{1}$) let

$$\mathcal{F}_{L+1} \equiv \mathcal{F}(\tilde{V} \backslash S_{L+1}) \equiv \sigma(\xi(\tilde{V} \backslash S_{L+1}))$$

denote the sigma-field generated by each of the vertices in $\partial a_L \cup \partial a$ after time 0 or its last appearance in $\tilde{\gamma}_{L+1}$ before t. Recall that $\mathsf{N}_L = \{\tilde{\gamma}_L \in \tilde{\Gamma}\}$. Let

$$\mathsf{N}_L^\circ \equiv \{Y_{t_\ell}(\partial a_\ell) = \underline{1}, \text{ for all } \ell \in [L] \backslash I_\star\} \supseteq \mathsf{N}_L.$$

By the Markov property of process Y_t, the event M'_{L+1}, defined by substituting the event C_{L+1} in the definition of $\mathsf{M}^{\mathrm{m}}_{L+1}$ with C'_{L+1} from (22), is independent of $\mathcal{F}_{L+1}$ given N_L°.

Meanwhile, for each $\ell \in I_{\mathrm{m}}$, the first five events in the definition of $\mathsf{M}^{\mathrm{m}}_\ell$ are measurable w.r.t. $\mathcal{F}_{L+1}$ while E_ℓ might depend also on the updates of ξ in $\partial a_{\ell-1} \times [t_{\ell-1}, t_\ell]$. In particular, E_ℓ is not $\mathcal{F}_{L+1}$-measurable if and only if

$$\ell \in I_{\mathsf{E}} \equiv \big\{\ell \in I_{\mathrm{m}} : \mathsf{E}_\ell \neq \varnothing^c \text{ and } \exists u \in (\partial a \cap \partial a_L^c) \backslash \{v\}, \ell - 1 = \ell_-(u)\big\}. \tag{26}$$

For each $\ell \in I_2$, the events $\mathsf{A}_\ell, \mathsf{C}_\ell$ and $\tilde{\mathsf{D}}_\ell$ are measurable w.r.t. $\mathcal{F}_{L+1}$ while $\tilde{B}_\ell$ might depend on the updates of ξ in $\partial a_{\ell-2} \times [t_{\ell-2}, t_\ell]$. More specifically, $\tilde{B}_\ell$ is not $\mathcal{F}_{L+1}$-measurable if and only if

$$\ell \in I_{\mathsf{B}} \equiv \big\{\ell \in I_2 : \exists u \in (\partial a \cap \partial a_L^c) \backslash \{v\}, \ell - 2 = \ell_-(u)\big\}. \tag{27}$$

Let

$$\mathsf{M}_\ell^{\mathrm{m,F}} \equiv \begin{cases} \mathsf{M}^{\mathrm{m}}_\ell & \ell \in I_{\mathrm{m}} \backslash I_{\mathsf{E}}, \\ \mathsf{A}_\ell \cap \mathsf{B}_\ell \cap \mathsf{C}_\ell \cap \mathsf{D}^1_\ell \cap \mathsf{D}^2_\ell & \ell \in I_{\mathsf{E}} \end{cases}, \quad \mathsf{M}_\ell^{\mathrm{II,F}} \equiv \begin{cases} \mathsf{M}^{\mathrm{II}}_\ell & \ell \in I_2 \backslash I_{\mathsf{B}}, \\ \mathsf{A}_\ell \cap \mathsf{C}_\ell \cap \tilde{\mathsf{D}}_\ell & \ell \in I_{\mathsf{B}}, \end{cases}.$$

be the $\mathcal{F}_{L+1}$-measurable part of $\mathsf{M}^{\mathrm{m}}_\ell$ and $\mathsf{M}^{\mathrm{II}}_\ell$ and let $\mathsf{N}_L^{\mathrm{F}} \equiv (\cap_{\ell \in I_{\mathrm{m}}} \mathsf{M}_\ell^{\mathrm{m,F}}) \cap (\cap_{\ell \in I_2} \mathsf{M}_\ell^{\mathrm{II,F}})$. We have

$$\mathbb{P}(\mathsf{M}^{\mathrm{m}}_{L+1} \mid \mathsf{N}_L, \mathsf{A}_{L+1}) \leqslant \mathbb{P}(\mathsf{M}'_{L+1} \mid \mathsf{N}_L, \mathsf{A}_{L+1}) = \mathbb{P}(\mathsf{M}'_{L+1} \mid \mathsf{N}_L^{\mathrm{F}}, \mathsf{A}_{L+1}, \cap_{\ell \in I_{\mathsf{E}}} \mathsf{E}_\ell, \cap_{\ell \in I_{\mathsf{B}}} \mathsf{B}_\ell).$$

To further simplify the conditioning part of the probability, we partition the events $\{\mathsf{E}_\ell\}_{\ell \in I_{\mathsf{E}}}$ and $\{\tilde{\mathsf{B}}_\ell\}_{\ell \in I_{\mathsf{B}}}$ into subsets such that each subset can be represented as the intersection of some $\mathcal{F}_{L+1}$-measurable event and $\mathcal{F}_{L+1}$-conditionally-independent event:

1. For each $\ell \in I_{\mathsf{E}}$, we split $\partial a_{\ell-1} \cap \partial a_\ell^c$ into the non-intersecting union of
$$W_\ell \equiv \{u \in \partial a_{\ell-1} \cap \partial a_\ell^c : u \in (\partial a \cap \partial a_L^c) \backslash \{v\}, \ell_-(u) = \ell - 1\}$$
and $V_\ell \equiv (\partial a_{\ell-1} \cap \partial a_\ell^c) \backslash W_\ell$. It follows from the definition of V_ℓ that $Y_{t_\ell}(V_\ell)$ is $\mathcal{F}_{L+1}$-measurable and $Y_{t_\ell}(W_\ell)$ is independent of $\mathcal{F}_{L+1}$ conditioned on N_L°. Let $\mathsf{E}^0_\ell \equiv \{Y_{t_\ell}(V_\ell) \neq \underline{1}\}$ and $\mathsf{E}^1_\ell \equiv \{Y_{t_\ell}(W_\ell) \neq \underline{1}\}$. Then E_ℓ can be partitioned into the events E^0_ℓ and $\mathsf{E}^1_\ell \backslash \mathsf{E}^0_\ell = \mathsf{E}^1 \cap (\mathsf{E}^0)^c$.
2. For each $\ell \in I_{\mathsf{B}}$, we similarly split $\partial a_{\ell-2} \cup \partial a_\ell$ into the non-intersecting union of
$$W_\ell \equiv \{u \in \partial a_{\ell-2} \cap \partial a_\ell^c : u \in (\partial a \cap \partial a_L^c) \backslash \{v\}, \ell_-(u) = \ell - 2\}$$
and $V_\ell \equiv (\partial a_{\ell-2} \cap \partial a_\ell^c) \backslash W_\ell$, and define $\tilde{W}_\ell \equiv W_\ell \times (t_{\ell-2}, t_\ell)$, $\tilde{V}_\ell \equiv V_\ell \times (t_{\ell-2}, t_\ell)$. Recall that ξ° is the unmarked update sequence, i.e., $(v, t) \in \xi^\circ$ if and only if $(v, t, 1) \in \xi$ or $(v, t, 0) \in \xi$. The event $\tilde{\mathsf{B}}_\ell$ is measurable w.r.t. the sigma field generated by $\xi^\circ(\tilde{W}_\ell \cup \tilde{V}_\ell) = \xi^\circ(\tilde{W}_\ell) \times \xi^\circ(\tilde{V}_\ell)$. More specifically, let $\Xi_\ell(\tilde{W}_\ell)$ be the set of possible configurations of $\xi^\circ(\tilde{W}_\ell)$ in event $\tilde{\mathsf{B}}_\ell$. Then $\tilde{\mathsf{B}}_\ell$ can be written as
$$\tilde{\mathsf{B}}_\ell = \cup_{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell)} \{\xi^\circ(\tilde{W}_\ell)\} \times \{\xi^\circ(\tilde{V}_\ell) : \xi^\circ(\tilde{W}_\ell \cup \tilde{V}_\ell) \in \tilde{\mathsf{B}}_\ell\},$$

where $\xi^\circ(\tilde{W}_\ell)$ is independent of $\mathcal{F}_{L+1}$ and $\xi^\circ(\tilde{V}_\ell)$ is $\mathcal{F}_{L+1}$-measurable.

By the fundamental formula of total probability, for any two events A, B and partition of A into disjoint sets $A = \cup_{i=1}^n A_i$, we have

$$\mathbb{P}(B \mid A) = \sum_{i=1}^n \mathbb{P}(B \mid A_i)\mathbb{P}(A_i \mid A) \leqslant \sup_{1 \leqslant i \leqslant n} \mathbb{P}(B \mid A_i).$$

Applying the same argument to the aforementioned partitions of E_ℓ and $\tilde{\mathsf{B}}_\ell$, we have

$$\begin{aligned}
&\mathbb{P}(\mathsf{M}'_{L+1} \mid \mathsf{N}^{\mathrm{F}}_L, \mathsf{A}_{L+1}, \cap_{\ell \in I_\mathsf{E}} \mathsf{E}_\ell, \cap_{\ell \in I_\mathsf{B}} \mathsf{B}_\ell) \\
&\leqslant \sup_{\substack{\xi^\circ(\tilde{W}_\ell \cup \tilde{V}_\ell) \in \tilde{\mathsf{B}}_\ell, \\ \ell \in I_\mathsf{B}; I'_\mathsf{E} \subseteq I_\mathsf{E}}} \mathbb{P}\left(\mathsf{M}'_{L+1} \,\middle|\, \mathsf{N}^{\mathrm{F}}_L, \mathsf{A}_{L+1}, \bigcap_{\ell \in I'_\mathsf{E}} [\mathsf{E}^1_\ell \cap (\mathsf{E}^0_\ell)^c], \bigcap_{\ell \in I_\mathsf{E} \setminus I'_\mathsf{E}} \mathsf{E}^0_\ell, \{\xi^\circ(\tilde{W}_\ell), \xi^\circ(\tilde{V}_\ell)\}_{\ell \in I_\mathsf{B}} \right) \\
&= \sup_{\substack{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell), \\ \ell \in I_\mathsf{B}; I'_\mathsf{E} \subseteq I_\mathsf{E}}} \mathbb{P}(\mathsf{M}'_{L+1} \mid \mathsf{N}^\circ_L, \mathsf{A}_{L+1}, \cap_{\ell \in I'_\mathsf{E}} \mathsf{E}^1_\ell, \{\xi^\circ(\tilde{W}_\ell)\}_{\ell \in I_\mathsf{B}}) \\
&\leqslant \sup_{\substack{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell), \\ \ell \in I_\mathsf{B}; I'_\mathsf{E} \subseteq I_\mathsf{E}}} \mathbb{P}(\mathsf{D}^1_{L+1} \mid \mathsf{N}^\circ_L, \mathsf{C}'_{L+1}, \cap_{\ell \in I'_\mathsf{E}} \mathsf{E}^1_\ell, \{\xi^\circ(\tilde{W}_\ell)\}_{\ell \in I_\mathsf{B}}) \cdot \mathbb{P}(\mathsf{B}_{L+1}, \mathsf{D}^2_{L+1}, \mathsf{E}_{L+1} \mid \mathsf{N}^\circ_L),
\end{aligned} \tag{28}$$

where the penultimate step uses the conditional independency of M'_{L+1} and $\mathcal{F}_{L+1}$ given N°_L, and the last step uses the independence of updates on $\partial a_L \times (t_L, \infty)$ and $\cup_{u \in \partial a \setminus \partial a_L} \{u\} \times (t_-(u), \infty)$.

To conclude the proof, we show that the first probability in (28) is uniformly bounded by $2^{-|\partial a \cap \partial a_L^c \setminus \{v\}|}$. Recall the definition of W_ℓ for each $\ell \in I_\mathsf{E} \cup I_\mathsf{B}$. For any $I'_\mathsf{E} \subseteq I_\mathsf{E}$ and $\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell), \ell \in I_\mathsf{B}$, we define

$$W_0 \equiv W_0(I'_\mathsf{E}) \equiv (\partial a \cap \partial a_L^c \setminus \{v\}) \setminus (\cup_{\ell \in I'_\mathsf{E} \cup I_\mathsf{B}} W_\ell).$$

The events $\{W_\ell\}_{\ell \in I'_\mathsf{E} \cup I_\mathsf{B} \cup \{0\}}$ form a partition of set $\partial a \cap \partial a_L^c \setminus \{v\}$. It follows that

$$\begin{aligned}
&\mathbb{P}(\mathsf{D}^1_{L+1} \mid \mathsf{N}^\circ_L, \mathsf{C}'_{L+1}, \cap_{\ell \in I'_\mathsf{E}} \mathsf{E}^1_\ell, \{\xi^\circ(\tilde{W}_\ell)\}_{\ell \in I_\mathsf{B}}) \\
&\quad = \prod_{\ell \in I'_\mathsf{E}} \mathbb{P}(Y_t(W_\ell) = \underline{1} \mid Y_{t_{\ell-1}}(W_\ell) = \underline{1}, Y_{t_\ell}(W_\ell) \neq \underline{1}, T_+(a; a_{\ell-1}, t_{\ell-1}) < t) \\
&\quad \cdot \prod_{\ell \in I_\mathsf{B}} \mathbb{P}(Y_t(W_\ell) = \underline{1} \mid Y_{t_{\ell-2}}(W_\ell) = \underline{1}, \xi^\circ(\tilde{W}_\ell), T_+(a; a_{\ell-2}, t_{\ell-2}) < t) \\
&\quad \cdot \prod_{u \in W_0(I'_\mathsf{E})} \mathbb{P}(Y_t(u) = \underline{1} \mid Y_{t_-(u)}(u) = \underline{1}, T_+(u; t_-(u)) < t).
\end{aligned}$$

For each probability in the first product, the monotonicity of the process Y_t implies that removing the condition of $Y_t(W_\ell) \neq \underline{1}$ will only increase its value. Thus

$$\mathbb{P}(\mathsf{D}^1_{L+1} \mid \mathsf{N}^\circ_L, \mathsf{C}'_{L+1}, \cap_{\ell \in I'_\mathsf{E}} \mathsf{E}^1_\ell, \{\xi^\circ(\tilde{U}_\ell)\}_{\ell \in I_\mathsf{B}}) \leqslant \prod_{\ell \in I'_\mathsf{E} \cup I_\mathsf{B} \cup \{0\}} 2^{-|U_\ell|} = 2^{-|\partial a \cap \partial a_L^c \setminus \{v\}|}.$$

Plugging the last equation back into (28) concludes the proof. □

Proof of Lemma 4.16. Denote $\underline{1} \equiv (1, 1, \ldots, 1)$ and $z = (0, 1, 1, \ldots, 1)$. We first explain how the case $Z_0 = \underline{1}$ implies all other cases:

For each $v \in \{0, 1\}^m$, let $\mathcal{E}_v \equiv \mathbb{E}_v[|\{0 \leqslant t \leqslant T_+ : Z_t = \underline{1}\}|]$ be the duration of time that the process Z_t, starting from v, stays at state $\underline{1}$ before T_+. By symmetry and monotonicity,

z achieves the maximum of $\mathcal{E}_v$ over $v \in \{0,1\}^m \backslash \{\underline{1}\}$ (along with other maximizers). Let T'_+ be the first time by which every coordinate, apart perhaps from the first one, is updated. Denote $\mathcal{E} \equiv \mathcal{E}_{\underline{1}}$. Then by first step analysis and symmetry

$$\mathcal{E} = 1 + \frac{1}{2}\mathcal{E}_z + \frac{1}{2}\mathcal{E}', \quad \text{where} \quad \mathcal{E}' \equiv \mathbb{E}_{\underline{1}}[|\{0 \leqslant t \leqslant T'_+ : Z_t = \underline{1}\}|].$$

Note that $\{T'_+ \neq T_+\}$ is precisely the event that the first coordinate is the last one to be updated. Given $T'_+ \neq T_+$ we have that $T'_+ - T_+$ has a Geometric$(1/m)$ distribution and that at each step t between T_+ and T'_+ the probability that $Z_t = \underline{1}$ is $2^{-(m-1)}$. Thus

$$\mathcal{E} = \mathcal{E}' + \mathbb{P}_{\underline{1}}[\{T'_+ \neq T_+\}]m2^{-(m-1)} = \mathcal{E}' + 2^{-(m-1)}.$$

Plugging this identity above yields that $\mathcal{E}_z = \mathcal{E} - 2 + 2^{-(m-1)}$.

We now treat the case $Z_0 = (1,1,\ldots,1)$. Note that after precisely i coordinates have already been updated, the probability that the chain is at $\underline{1}$ is 2^{-i}. The number of such steps follows a Geometric distribution with parameter $(m-i)/m$. Thus the desired expectation is

$$\mathcal{E} = \sum_{i=0}^{m-1} \frac{2^{-i}m}{m-i} = m2^{-m}\sum_{i=1}^{m}\frac{2^i}{i}.$$

We now proceed to give an upper-bound on $\mathcal{E}$. Let $(c_i)_{i\in\mathbb{Z}_+}$ be a sequence of real numbers and denote $d_i \equiv 2^i c_i$. Recall that by Abel's summation by parts formula, using the fact that $2^{i+1} - 2^i = 2^i$, we get that for any integers $n_2 \geqslant n_1 \geqslant 0$,

$$\sum_{i=n_1}^{n_2} d_i = (d_{n_1} - d_{n_1+1}) + 2d_{n_2} + \sum_{i=n_1+1}^{n_2-1} 2^{i+1}(c_i - c_{i+1}).$$

Applying the Abel's summation formula repeatedly (and noting that at each iteration the first and second term cancel out) yields that

$$\begin{aligned}
\sum_{i=1}^{m}\frac{2^i}{i} &= \frac{2^{m+1}}{m} + \sum_{i=2}^{m-1}\frac{2^{i+1}}{i(i+1)} = \frac{2^{m+1}}{m} + \frac{2^{m+1}}{m(m-1)} + \sum_{i=3}^{m-2}\frac{2^{i+2}(2!)}{i(i+1)(i+2)} \\
&= \cdots = \frac{2^m}{m}\sum_{j=0}^{\lceil m/2\rceil-1}\frac{2}{\binom{m-1}{j}} + \sum_{i=\lceil m/2\rceil}^{\lfloor m/2\rfloor+1}\frac{2^{i+\lceil m/2\rceil-1}[(\lceil m/2\rceil-1)!]}{i(i+1)\ldots(i+\lceil m/2\rceil-1)} \\
&\leqslant \frac{2^m}{m}\left(2^{-\lceil m/2\rceil+1_{m\in2\mathbb{N}}} + \sum_{j=0}^{\lceil m/2\rceil-1} 2\binom{m-1}{j}^{-1}\right).
\end{aligned}$$

This yields that $\mathcal{E} \leqslant 2\big(\sum_{j=0}^{\lceil m/2\rceil-1}\binom{m-1}{j}^{-1}\big) + 2^{-\lceil m/2\rceil+1_{m\in2\mathbb{N}}}$. Checking each case separately, it is not hard to to verify that for $m < 7$ we have that $\mathcal{E} \leqslant 2 + \frac{6}{m} - 2^{-(m-1)}$, whereas if $m \geqslant 7$ we have that $\mathcal{E} \leqslant 2 + \frac{2}{m-1} + \frac{10}{(m-1)(m-2)} + 2^{-\lceil m/2\rceil+1_{m\in2\mathbb{N}}} \leqslant 2 + \frac{6}{m} - 2^{-(m-1)}$. □

Proof of Lemma 4.17. Recall that for $v \in V$, $b \in F$ and $s \geqslant 0$, we have that $T_+(v; b, s) = s + [T_+(v; s) - s] \cdot \mathbf{1}\{v \in \partial b\}$. Similarly to the construction of (22), we define

$$\mathsf{C}'_{L+2}(t) \equiv \bigcap_{u \in \partial a \setminus \partial a_L} \Big[\{T_+(u; 0) \leqslant t\} \cap \big(\cap_{\ell=1}^{L} \{T_+(u; a_\ell, t_\ell) \leqslant t\} \big) \Big],$$

$$\mathsf{C}''_{L+2}(t) \equiv \bigcap_{u \in \partial a \cap \partial a_L \setminus \{v\}} \big(\cap_{\ell=1}^{L} \{T_+(u; a_\ell, t_\ell) \leqslant t\} \big)$$

and

$$\tilde{\mathsf{D}}^1_{L+2}(t) \equiv \{Y_t(\partial a \cap \partial a_L^c \setminus \{v\}) = \underline{1}\}, \quad \tilde{\mathsf{D}}^2_{L+2}(t) \equiv \{Y_t(\partial a \cap \partial a_L \setminus \{v\}) = \underline{1}\}.$$

For every $u \in \partial a \cap \partial a_L$, the event $\{T_+(u; 0) \leqslant t\}$ is implied by $\{T_+(u; a_L, t_L) \leqslant t\}$. In particular, we have that $\mathsf{C}_{L+2}(t) \subseteq \mathsf{C}'_{L+2}(t) \cap \mathsf{C}''_{L+2}(t)$. Fix the choice of t and suppress it from the notation. Applying similar reasoning as in the proof of Lemma 4.15 (and using the notation from that proof), we get that

$$\begin{aligned}
&\mathbb{P}(\mathsf{M}^{\mathrm{II}}_{L+2} \mid \mathsf{N}_L, \mathsf{A}_{L+1}) \\
&\leqslant \sup_{\substack{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell), \\ \ell \in I_{\mathsf{B}}; I'_{\mathsf{E}} \subseteq I_{\mathsf{E}}}} \mathbb{P}(\tilde{\mathsf{B}}_{L+2}, \mathsf{C}_{L+2}, \tilde{\mathsf{D}}^1_{L+2}, \tilde{\mathsf{D}}^2_{L+2} \mid \mathsf{N}^\circ_L, \cap_{\ell \in I'_{\mathsf{E}}} \mathsf{E}^1_\ell, \{\xi^\circ(\tilde{W}_\ell)\}_{\ell \in I_{\mathsf{B}}}) \\
&\leqslant \mathbb{P}(\tilde{\mathsf{D}}^1_{L+2} \mid \tilde{\mathsf{B}}_{L+2}, \mathsf{C}'_{L+2}) \cdot \mathbb{P}(\tilde{\mathsf{D}}^2_{L+2} \mid \tilde{\mathsf{B}}_{L+2}, \mathsf{C}''_{L+2}) \cdot \mathbb{P}(\tilde{\mathsf{B}}_{L+2}) = 2^{-(k-1)} \mathbb{P}(\tilde{\mathsf{B}}_{L+2}).
\end{aligned}$$

□

5. Random regular hypergraph

In this section we exploit the locally tree-like geometry of random regular hypergraphs and prove Theorem 1.3. For two vertices $v, v' \in V$, we define the distance $d(v, v')$ as the number of hyperedges on the shortest path from v to v'. The property we will need is the following.

Definition 5.1. For each $R \geqslant 1$, we say that G is R-good if for every $v \in V$, the R-neighbourhood of v (as a subgraph of the bipartite graph representation of G) contains at most one cycle.

A similar argument as the proof for random regular graph (i.e., $k = 2$) in [8] yields the following.

Proposition 5.2 (cf. [8] Lemma 2.1)**.** *For any constant $R \geqslant 0$ and $G \sim \mathcal{H}(n, d, k)$.*

$$\lim_{n \to \infty} \mathbb{P}(G \text{ is } R\text{-good}) = 1.$$

In the remainder of the section we are going to fix $R_\star = R_\star(\Delta, k) \geqslant 2$ to be a constant to be determined later and restrict our attention to the following subset of n-vertex hypergraphs:

$$\mathcal{G} \equiv \mathcal{G}_n(R_\star) \equiv \{G : G \text{ is } \Delta\text{-regular, } k\text{-uniform and } R_\star\text{-good}\}.$$

5.1. **Projected path.**

We define for every $a \in F$ the subset

$$\mathsf{cyc}(a) \equiv \{v \in \partial a : v \text{ is contained in a cycle shorter than } 2R_\star\} \subseteq \partial a,$$

where the length of cycle is the number of vertices along the cycle. For each $G \in \mathcal{G}$, the definition of $R_\star$-good implies that $|\mathsf{cyc}(a)| \leqslant 2$ for all $a \in F$. In particular, for every $a \in F$ there is at most one $a' \in F$ such that $|\partial a \cap \partial a'| \geqslant 2$, in which case $\mathsf{cyc}(a) = \mathsf{cyc}(a') = \partial a \cap \partial a'$.

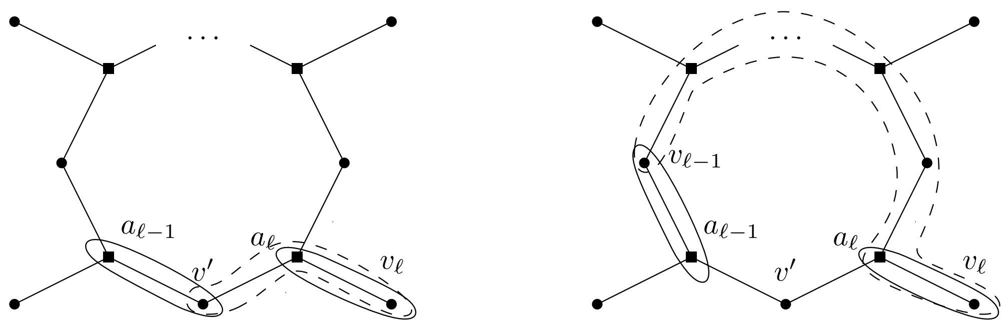


FIGURE 3. Direct step and cycle step

For each discrepancy sequence $\zeta = ((u_i, b_i, s_i))_{0\leqslant i\leqslant M}$, let $\gamma_L \equiv \gamma_{L(\zeta)}(\zeta) \equiv ((v_\ell, a_\ell, t_\ell))_{0\leqslant \ell\leqslant L}$ be the minimal path constructed according to the proof of Lemma 4.5 and for each $1 \leqslant \ell \leqslant L$, let $i(\ell)$ be the step it corresponds to (via (14)). From the construction of Lemma 4.5, we can observe that for each (v_ℓ, a_ℓ, t_ℓ), $1 \leqslant \ell \leqslant L$, there exists an alternating sequence of vertices and hyperedges and an increasing subsequence of indices

$$(\tilde{a}_0\tilde{v}_0\tilde{a}_1\tilde{v}_1\cdots\tilde{a}_m\tilde{v}_m), \quad i(\ell-1) = j_0 < j_1 < \cdots < j_m = i(\ell)$$

that "represents" a subsequence in ζ, i.e., it satisfies

$$\tilde{v}_r \in \partial\tilde{a}_r \cap \partial\tilde{a}_{r+1},\ 0 \leqslant r \leqslant m-1, \quad (\tilde{v}_r, \tilde{a}_r) = (u_{j_r}, b_{j_r}),\ 0 \leqslant r \leqslant m. \tag{29}$$

For each ℓ such that $a_\ell \neq a_{\ell-1}$ and $\partial a_\ell \cap \mathsf{cyc}(a_{\ell-1}) = \varnothing$, there must exist v' such that $\partial a_\ell \cap \partial a_{\ell-1} = \{v'\}$. In order for a_ℓ to remain "active" with respect to $a_{\ell-1}$ by time t_ℓ, namely $t_\ell \leqslant T_+(a_\ell; a_{\ell-1}, t_{\ell-1})$, there must not be any $1 \leqslant i \leqslant m$ such that $\tilde{v}_i = v'$. In particular, it implies that either $v_{\ell-1} = v'$ or $(\tilde{a}_0\tilde{v}_0\tilde{a}_1\tilde{v}_1\cdots\tilde{a}_m v')$ completes a cycle in G. For $v \in \partial a \in F$ we write $\mathsf{cyc}^+(a, v) \equiv \mathsf{cyc}(a) \cup \{v\}$. For each path $\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L} \in \Gamma_L$ we define (See Figure 3)

$$I_{\mathrm{c}} \equiv \{1 \leqslant \ell \leqslant L : \partial a_\ell \cap \mathsf{cyc}^+(a_{\ell-1}, v_{\ell-1}) = \varnothing\}$$

be the set of *cycle steps* and define $I_{\mathrm{d}} \equiv \{1, \ldots, L\} \backslash I_{\mathrm{c}}$ to be the set of *direct steps*. It follows that for given (v_ℓ, a_ℓ, t_ℓ), there is at most $\Delta \cdot |\{v_\ell\} \cup \mathsf{cyc}(a_\ell)| \leqslant 3\Delta$ ways to select $a_{\ell+1}$ such that the next step is a direct step.

Definition 5.3. For each path $\gamma_L \in \Gamma_L$, we say that γ_L is a *(relaxed) projected path* if for each $1 \leqslant \ell \in L$ it satisfies the five events $\mathsf{A}_\ell, \mathsf{B}_\ell, \mathsf{C}_\ell, \mathsf{D}^1_\ell, \mathsf{D}^2_\ell$ defined in (11) and for each $\ell \in I_{\mathrm{c}}$, it further satisfies

$$\mathsf{G}_\ell \equiv \left\{\begin{array}{l} \exists \text{ an alternating sequence } (\tilde{a}_0\tilde{v}_0\tilde{a}_1\tilde{v}_1\cdots\tilde{a}_m\tilde{v}_m) \text{ such that:} \\ 1.\ (\tilde{v}_0, \tilde{a}_0) = (v_{\ell-1}, a_{\ell-1}),\ (\tilde{v}_m, \tilde{a}_m) = (v_\ell, a_\ell) \\ 2.\ (\tilde{a}_0\tilde{v}_0\tilde{a}_1\tilde{v}_1\cdots\tilde{a}_m v'_\ell) \text{ completes a cycle in } G. \\ 3.\ \exists t_{\ell-1} = \tilde{t}_0 \leqslant \tilde{t}_1 < \cdots < \tilde{t}_{m-1} \leqslant \tilde{t}_m = t_\ell, \text{ such that} \\ \quad (\tilde{v}_r, \tilde{t}_r) \in \xi^\circ \text{ for } 1 \leqslant r \leqslant m-2. \end{array}\right\},$$

where v'_ℓ is the vertex in $\partial a_\ell \cap \partial a_{\ell-1}$. Let $\tilde{\Gamma}_{\mathrm{proj},L}$ denotes the set of (relaxed) projected path.

Remark 5.4. Recall the definition in Remark 4.6. The discussion preceding the definition implies that $\Gamma_{\mathrm{proj},L} \subseteq \tilde{\Gamma}_{\mathrm{proj},L}$. Meanwhile a path in $\tilde{\Gamma}_{\mathrm{proj},L}$ does not necessarily satisfy events $\{\mathsf{D}^3_\ell\}$. We also do not assume in the definition of G_ℓ that the values of updates at $\{(\tilde{v}_r, \tilde{t}_r)\}$ are all ones (which turns out to be crucial in the proof). Hence the name relaxed.

Remark 5.5. In the third requirement of the definition of G_ℓ, we require $(\tilde{v}_r, \tilde{t}_r) \in \xi^\circ$ for $1 \leqslant r \leqslant m-2$ so that none of the $\tilde{v}_r$'s belong to $\partial a_{\ell-1} \cup \partial a_\ell$ and hence G_ℓ and B_ℓ depend on different vertices. In the proof, we shall condition on events of the form G_ℓ. While conditioning on updates with value one works against us, conditioning on having at earlier times updates with unspecified values can only work to our advantage.

We now outline the two-step recursion on $\tilde{\Gamma}_{\mathrm{proj},L}$ and the proof of Theorem 1.3. Most of the arguments are parallel to the corresponding parts in Section 4. With some abuse of notation, we occasionally override the notations in Section 4 with slightly different meanings.

Fix $\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant \ell \leqslant L} \in \Gamma_L$ and vertex-hyperedge pair (v,a) satisfying $a \in \mathbb{N}(a_L), v \in \partial a$, we let $\gamma_{L+1}(t) \equiv \gamma_{L+1}(v,a,t)$ be the path extended from γ_L with $(v_{L+1}, a_{L+1}, t_{L+1}) = (v,a,t)$ and define

$$N^{\mathrm{proj}}(v,a) \equiv N^{\mathrm{proj}}(v,a;\gamma_L) \equiv |\{t : \gamma_{L+1}(v,a,t) \in \tilde{\Gamma}_{\mathrm{proj},L+1}\}|.$$

Lemma 5.6. *For any two neighbouring hyperedges $a \in F, b \in \mathbb{N}(a)$, let $R(a,b)$ be the length of the shortest cycle in G that contains a and b. There exists a constant $C_1 > 0$ such that for any integer $L \geqslant 1$, $\gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}$ and $a \in \mathbb{N}(a_L)$, $v \in \partial a$,*

$$\mathbb{E}[N^{\mathrm{proj}}(v,a) \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant C_1 \begin{cases} 1/k & a = a_L, \\ 2^{-k} & a \neq a_L, \partial a \cap \mathsf{cyc}^+(a_L, v_L) \neq \varnothing, \\ p_{\mathrm{c}}(R(a,a_L)) & a \neq a_L, \partial a \cap \mathsf{cyc}^+(a_L, v_L) = \varnothing, \end{cases} \tag{30}$$

where

$$p_{\mathrm{c}}(r) \equiv e^{-k} + k \sum_{m \geqslant r} (\Delta k)^m \mathbb{P}(\mathrm{Pois}(k) \geqslant m-2).$$

Proposition 5.7. *Denote $R_\star = R_\star(\Delta,k) \equiv \lceil e\Delta k^2 \rceil + 1$. Then that for all $\Delta \leqslant 2^k$, $\gamma_L \in \Gamma_L$, $r \geqslant 2R_\star$ and $1 \leqslant \ell \leqslant L$,*

$$p_c(r) \leqslant (1 + 2^{-e\Delta k^2}) \cdot e^{-k} \leqslant 2e^{-k}.$$

Proof. Let R follows the Poisson(k) distribution of parameter k. Using Stirling's approximation, we have that $\mathbb{P}(R=m) \leqslant \frac{e^{-k}k^m}{m!} \leqslant \frac{e^{-k}}{\sqrt{2\pi m}} \left(\frac{ek}{m}\right)^m$. Hence $\mathbb{P}(R \geqslant m) \leqslant 2\mathbb{P}(R=m)$ for all $m \geqslant 2ek$. It follows that for all $r \geqslant \lceil 2e\Delta k^2 \rceil + 2$,

$$\begin{aligned} p_c(r) - e^{-k} &= k \sum_{m \geqslant r} (\Delta k)^m \mathbb{P}(R \geqslant m-2) \leqslant 2\Delta k^2 \sum_{m \geqslant r} \frac{e^{-k}}{\sqrt{2\pi(m-2)}} \left(\frac{e\Delta k^2}{m}\right)^{m-2} \\ &\leqslant \frac{1}{2}\Delta k e^{-k} \sum_{i \geqslant r-2} 2^{-i} \leqslant \Delta k e^{-k} 2^{-(r-2)} \leqslant e^{-k} 2^{-e\Delta k^2}. \end{aligned}$$

This concludes the proof. □

Meanwhile, similarly to the definition of type-II branchings in redacted paths, special treatment is needed for paths staying at the same hyperedges three steps in a row. Fix $\gamma_L = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L} \in \Gamma_L$ and let vertices v, v' satisfy $v, v' \in a_L$. Let $\gamma_{L+2} \equiv \gamma_{L+2}(v,v',t,t') = ((v_\ell, a_\ell, t_\ell))_{0\leqslant\ell\leqslant L+2} \in \Gamma_{L+2}$ be the path extended from γ_L with

$$(v_{L+2}, a_{L+2}, t_{L+2}) = (v, a_L, t), \quad (v_{L+1}, a_{L+1}, t_{L+1}) = (v', a_L, t'),$$

Observe that by construction, in order for γ_{L+2} (as above) to be in $\tilde{\Gamma}_{\mathrm{proj},L+2}$, it must be the case that a_L is deactivated w.r.t. itself by time t (i.e., $t > T_+(a_L;t_L)$), making the event

$Y_t(\partial a_L) = \underline{1}$ unlikely (cf. the first case of Lemma 5.6). This is quantified in the following lemma. We define

$$N^{\mathrm{proj}}_{=}(v,v') \equiv N^{\mathrm{proj}}_{=}(v,v';\gamma_L) \equiv |\{(t,t') : \gamma_{L+2}(v,v',t,t') \in \tilde{\Gamma}_{\mathrm{proj},L+2}\}|.$$

Lemma 5.8. *Under the notations above, there exists an absolute constant $C_2 > 0$ such that*

$$\mathbb{E}[N^{\mathrm{proj}}_{=}(v,v';\gamma_L) \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant C_2 k^2 2^{-k}. \tag{31}$$

For $r = 1, 2$, Let $N_{\mathrm{proj},L+r}(\gamma_L)$ be the number of paths $\gamma_{L+r} \in \tilde{\Gamma}_{\mathrm{proj},L+r}$ that agree with γ_L in the first L steps and write $N^{\neq}_{\mathrm{proj},L+1}(\gamma_L)$ (resp. $N^{=}_{\mathrm{proj},L+1}(\gamma_L)$) for the number of paths γ_{L+1}'s counted in $N_{\mathrm{proj},L+1}(\gamma_L)$ that further satisfies $a_{L+1} \neq a_L$ (resp. $a_{L+1} = a_L$). Lemma 5.6, Lemma 5.8 and Proposition 5.7 together imply the following theorem. The proof is presented for the sake of completeness.

Theorem 5.9. *Under the above notation, there exists an absolute constant $C_{\mathrm{proj}} > 0$, such that for any $R_\star$-good hypergraph G, integer $L \geqslant 1$ and path $\gamma_L \in \Gamma_L$, we have that*

$$\mathbb{E}[N^{=}_{\mathrm{proj},L+1}(\gamma_L) \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant C_{\mathrm{proj}}, \tag{32}$$

$$\mathbb{E}[N^{\neq}_{\mathrm{proj},L+1}(\gamma_L) \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant C_{\mathrm{proj}}(\Delta k)2^{-k}, \tag{33}$$

$$\mathbb{E}[N_{\mathrm{proj},L+2}(\gamma_L) \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant C^2_{\mathrm{proj}}[(\Delta k 2^{-k}) + k^4 2^{-k}]. \tag{34}$$

Proof. Fix $\gamma_L \equiv ((v_\ell, a_\ell, t_\ell))_{1\leqslant \ell \leqslant L} \in \Gamma_L$ and for brevity define $\tilde{\mathbb{E}}_L \equiv \mathbb{E}[\,\cdot \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}]$. We first prove (32). By the first case of Lemma 5.6,

$$\tilde{\mathbb{E}}_L[N^{=}_{\mathrm{proj},L+1}(\gamma_L)] = \sum_{v \in \partial a_L} \tilde{\mathbb{E}}_L[N^{\mathrm{proj}}(v, a_L;\gamma_L)] \leqslant k \cdot C_1 k^{-1}.$$

We now prove (33), we define

$$A_{L+1} \equiv \{a \in \mathbb{N}(a_L)\backslash\{a_L\} : \partial a \cap \mathsf{cyc}^+(a_L, v_L) = \varnothing\}$$

be the set of hyperedges that could form a direct branching from (v_L, a_L, t_L). It follows that

$$N^{\neq}_{\mathrm{proj},L+1}(\gamma_L) = \sum_{a \in A_{L+1}} \sum_{v \in \partial a} N^{\mathrm{proj}}(v,a;\gamma_L) + \sum_{a \in \mathbb{N}(a_L)\backslash(A_{L+1}\cup\{a_L\})} \sum_{v\in\partial a} N^{\mathrm{proj}}(v,a;\gamma_L).$$

Observe that every $a \in \mathbb{N}(a_L)\backslash(A_{L+1} \cup \{a_L\})$ must satisfy $R(a, a_L) \geqslant R_\star$. Applying the last two cases of Lemma 5.6 then yields

$$\begin{aligned}\tilde{\mathbb{E}}_L[N^{\neq}_{\mathrm{proj},L+1}(\gamma_L)] &\leqslant k|A_{L+1}| \cdot C_1 2^{-k} + k|\mathbb{N}(a_L)| \cdot C_1 p_c(R_\star) \\ &\leqslant C_1[3k\Delta 2^{-k} + k\Delta \cdot k p_c(R^\star)] \leqslant C_{\mathrm{proj}}\,\Delta k 2^{-k}.\end{aligned}$$

Finally, to prove (34), we note that the bound of Lemma 5.6 does not depend on t_L. Again we abbreviate $\tilde{\mathbb{E}}_{L+1}(\cdot) \equiv \mathbb{E}[\cdot \mid \gamma_{L+1} \in \tilde{\Gamma}_{\mathrm{proj},L+1}]$ and let A_{L+1} be defined as in the previous case (with respect to the $(L+1)$'th step). There are three possible ways of extending γ_L to γ_{L+2}:

$$a_L = a_{L+1} = a_{L+2} \quad \text{or} \quad a_L = a_{L+1} \neq a_{L+2} \quad \text{or} \quad a_L \neq a_{L+1}.$$

Following a similar argument as that of Theorem 4.14, we have that

$$\begin{aligned}\tilde{\mathbb{E}}[N_{\mathrm{proj},L+2}(\gamma_L)] \leqslant & \sum_{v,v'\in a_L} \tilde{\mathbb{E}}_L[N^{\mathrm{proj}}_{=}(v,v')] \\ & + \sum_{a\neq a_L}\sum_{v\in\partial a} \tilde{\mathbb{E}}_L[N^{\mathrm{proj}}(v,a;\gamma_L)]\cdot \tilde{\mathbb{E}}_{L+1}[N^{=}_{\mathrm{proj},L+2}(\gamma_{L+1}(v,a))] \\ & + \sum_{v\in\partial a_L} \tilde{\mathbb{E}}_L[N^{\mathrm{proj}}(v,a_L;\gamma_L)]\cdot \tilde{\mathbb{E}}_{L+1}[N^{\neq}_{\mathrm{proj},L+2}(\gamma_{L+1}(v,a_L))].\end{aligned}$$

Applying (32), (33) and Lemma 5.8 implies that

$$\begin{aligned}\tilde{\mathbb{E}}[N_{\mathrm{proj},L+2}(\gamma_L)] \leqslant C_2k^4 2^{-k} + \tilde{\mathbb{E}}_L[N^{\neq}_{\mathrm{proj},L+1}(\gamma_L)] \max_{v,a:a\neq a_L, v\in\partial a} \tilde{\mathbb{E}}_{L+1}[N^{=}_{\mathrm{proj},L+2}(\gamma_{L+1}(v,a))] \\ + \tilde{\mathbb{E}}_L[N^{=}_{\mathrm{proj},L+1}(\gamma_L)] \max_{v\in\partial a_L} \tilde{\mathbb{E}}_{L+1}[N^{\neq}_{\mathrm{proj},L+2}(\gamma_{L+1}(v,a_L))] \\ \leqslant \text{ RHS of } (34),\end{aligned}$$

which concludes the proof. $\square$

Proof of Theorem 1.3. The assertion of Theorem 1.3 is obtained by combining Lemma 4.5, (15), Remark 5.4 and (34). $\square$

5.2. **Proof of Lemma 5.6 and Lemma 5.8.** In this subsection we present the proofs of remaining lemmas. We begin with Lemma 5.6. Throughout the proof, we keep $\gamma_L \in \Gamma_L, a \in \mathbb{N}(a_L), v \in \partial a$ fixed, and for brevity of notation write $\gamma_{L+1}(t)$ for the path extended from γ_L with $(v_{L+1}, a_{L+1}, t_{L+1}) = (v, a, t)$. We further define

$$\begin{aligned}\mathsf{M}^{\mathrm{c}}_\ell(t) \equiv \mathsf{M}^{\mathrm{c}}_\ell(\tilde{\gamma}_{L+1}(t)) \equiv \mathsf{A}_\ell(t)\cap\mathsf{B}_\ell(t)\cap\mathsf{C}_\ell(t)\cap\mathsf{D}^1_\ell(t)\cap\mathsf{D}^2_\ell(t)\cap\mathsf{G}_\ell(t) \quad \forall \ell\in I_{\mathrm{c}}(\gamma_{L+1}(t)), \\ \mathsf{M}^{\mathrm{d}}_\ell(t) \equiv \mathsf{M}^{\mathrm{d}}_\ell(\tilde{\gamma}_{L+1}(t)) \equiv \mathsf{A}_\ell(t)\cap\mathsf{B}_\ell(t)\cap\mathsf{C}_\ell(t)\cap\mathsf{D}^1_\ell(t)\cap\mathsf{D}^2_\ell(t) \quad \forall \ell\in I_{\mathrm{d}}(\gamma_{L+1}(t)),\end{aligned}$$

where we henceforth omit t from the notation for all events with $\ell \leqslant L$, as they do not depend on the value of t. We further write (overriding any conflicting definitions from Section 4)

$$\mathsf{N}_L \equiv \{\gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}\} = \Big[\cap_{\ell\in I_{\mathrm{c}}} \mathsf{M}^{\mathrm{c}}_\ell\Big] \bigcap \Big[\cap_{\ell\in I_{\mathrm{d}}} \mathsf{M}^{\mathrm{d}}_\ell\Big].$$

By Campbell's theorem,

$$\mathbb{E}[N^{\mathrm{proj}}(v,a)\,|\,\gamma_L\in\tilde{\Gamma}_{\mathrm{proj},L}] = \mathbb{E}\Big[\sum_{t:(v,t,1)\in\xi} \mathbf{1}\{\mathsf{M}^{\bullet}_{L+1}(t)\}\,\Big|\,\mathsf{N}_L\Big] = \frac{1}{2}\int_{t_L}^{\infty} \mathbb{P}(\mathsf{M}^{\bullet}_{L+1}(t)\,|\,\mathsf{N}_L, \mathsf{A}_{L+1}(t))dt.$$

where $\bullet = \mathrm{c}$ if $L+1 \in I_{\mathrm{c}}$ and $\bullet = \mathrm{d}$ otherwise.

Lemma 5.10. *Under the notations above,*

$$\mathbb{P}(\mathsf{M}^{\mathrm{d}}_{L+1}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t)) \leqslant \begin{cases} 2^{2-k}\mathbb{P}(\mathsf{B}_{L+1}(t)) & a\neq a_L \\ \mathbb{P}(\mathsf{B}_{L+1}(t)) & a = a_L\end{cases}.$$

Moreover if $L+1\in I_{\mathrm{c}}(\gamma_{L+1}(t))$, then

$$\mathbb{P}(\mathsf{M}^{\mathrm{c}}_{L+1}(t)\mid\mathsf{N}_L,\mathsf{A}_{L+1}(t)) \leqslant 2^{1-k}\mathbb{P}(\mathsf{G}_{L+1}(t))\mathbb{P}(\mathsf{B}_{L+1}(t)).$$

Proof. Recall the definition of $\ell_-(u), \mathcal{F}_{L+1}, \mathsf{N}_L^\circ, \mathsf{C}'_{L+1}$ from the proof of Lemma 4.15 and omit t from the notation. By the Markov property of the process Y_t, the events A_{L+1}, B_{L+1}, C'_{L+1}, D^1_{L+1}, D^2_{L+1} and (in the case of $L+1 \in I_\mathrm{c}$) G_{L+1} are independent of $\mathcal{F}_{L+1}$ given N_L°. Meanwhile, for each $1 \leqslant \ell \leqslant L$, we have that $\mathsf{M}^\mathrm{d}_\ell$ is measurable w.r.t. $\mathcal{F}_{L+1}$. It is left to treat $\{\mathsf{G}_\ell\}_{\ell \in I_\mathrm{c}}$.

Observe that for each $\ell \in I_\mathrm{c}$, we have that G_ℓ is a measurable function of $\xi^\circ(V \times [t_{\ell-1}, t_\ell))$, the time and locations of all updates between $t_{\ell-1}$ and t_ℓ without the value of the updates. Following a similar argument of Lemma 4.15, we define

$$W_\ell \equiv \{u \in \partial a_{\ell-1} \cap \partial a_\ell^c : u \in (\partial a \cap \partial a_L^c)\backslash\{v\}, \ell_-(u) = \ell - 1\}$$

and (overriding the definition in Lemma 4.15)

$$\tilde{W}_\ell \equiv W_\ell \times [t_{\ell-1}, t_\ell), \quad \tilde{V}_\ell \equiv W_\ell^c \times [t_{\ell-1}, t_\ell).$$

Let $\Xi_\ell(\tilde{W}_\ell)$ be the range of the unmarked update process $\xi^\circ(\tilde{W}_\ell)$ over G_ℓ (i.e., the collection of all possible values of $\xi^\circ(\tilde{W}_\ell)$, provided that G_ℓ occurs). It follows that

$$\mathsf{G}_\ell = \cup_{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell)} \{\xi^\circ(\tilde{W}_\ell)\} \times \{\xi^\circ(\tilde{V}_\ell) : \xi^\circ(\tilde{W}_\ell \cup \tilde{V}_\ell) \in \mathsf{G}_\ell\},$$

where $\xi^\circ(\tilde{W}_\ell)$ is independent of $\mathcal{F}_{L+1}$ and $\xi^\circ(\tilde{V}_\ell)$ is $\mathcal{F}_{L+1}$-measurable. Therefore following a similar calculation to (28), we have

$$\begin{aligned}
&\mathbb{P}(\mathsf{M}^\mathrm{d}_{L+1}(t) \mid \mathsf{N}_L, A_{L+1}(t)) \\
&\leqslant \sup_{\xi^\circ(\tilde{W}_\ell) \in \Xi_\ell(\tilde{W}_\ell)} \mathbb{P}(\mathsf{D}^1_{L+1} \mid \mathsf{N}_L^\circ, \mathsf{C}'_{L+1}, \{\xi^\circ(\tilde{W}_\ell)\}_{\ell \in I_\mathrm{c}}) \cdot \mathbb{P}(\mathsf{B}_{L+1}, \mathsf{D}^2_{L+1} \mid \mathsf{N}_L^\circ) \\
&\leqslant 2^{-(k-2)\mathbf{1}\{a \neq a_L\}} \mathbb{P}(\mathsf{B}_{L+1}),
\end{aligned} \tag{35}$$

where for the last step, we note that for $G \in \mathcal{G}$, we have that $|\partial a \cap \partial a_L| \leqslant 2$ if $a \neq a_L$.

Recall that the definition of G_{L+1} does not involve the updates on $\partial a_L \cup \partial a$. The result for $\mathsf{M}^\mathrm{c}_{L+1}$ follows a similar argument to that of (35). ☐

Proof of Lemma 5.6. The first two cases follow from a similar argument to that of Lemma 4.12 with overlap $m_{L+1} = k$ and $m_{L+1} \leqslant 2$, respectively. Here we only present the proof of the third case, leaving the first two as an exercise. Fix (v, a) such that $a \neq a_L$ and $\partial a \cap \mathsf{cyc}^+(a_L, v_L) = \varnothing$. Let $\gamma_{L+1} = \gamma_{L+1}(v, a; \gamma_L)$. We can write

$$\begin{aligned}
\int_{t_L}^\infty \mathbb{P}(\mathsf{G}_{L+1}(t))\mathbb{P}(\mathsf{B}_{L+1}(t))dt &\leqslant \int_{t_L}^{t_L+k} \mathbb{P}(\mathsf{G}_{L+1}(t))dt + \int_{t_L+k}^\infty \mathbb{P}(\mathsf{B}_{L+1}(t))dt. \\
&\leqslant k\mathbb{P}(\mathsf{G}_{L+1}(t_L+k)) + \mathbb{P}(T_+(a_{L+1}; a_L, t_L) > t_L + k)
\end{aligned} \tag{36}$$

Since in a cycle step $|\partial a_L \cap \partial a_{L+1}| = 1$, the second term on the RHS of (36) can be bounded by e^{-k}. For the first term, we enumerate over all possible cycles containing a_L, a_{L+1}. Fix some cycle $(\tilde{a}_0\tilde{v}_0\tilde{a}_1 \ldots \tilde{a}_m\tilde{v}_m)$ with $\tilde{a}_0 = a_L$ and $\tilde{a}_m = a_{L+1}$. Let $s_0 \equiv t_L$ and inductively define $s_i = T(\tilde{v}_i; s_{i-1})$, for all $1 \leqslant i \leqslant m-2$. Denote $\delta_i \equiv s_i - s_{i-1}$ and $S \equiv \sum_{i=1}^{m-2} \delta_i$. Then $\delta_1, \ldots, \delta_{m-2}$ are i.i.d. Exp(1) r.v.'s. In particular, interpreting the δ_i's as spacings between arrivals of a rate 1 Poisson process, we get that $\mathbb{P}(S \leqslant k) = \mathbb{P}(N_k \geqslant m-2)$, where N_k has a Poisson distribution of parameter k. In conclusion,

$$\mathbb{P}(\tilde{v}_1, \ldots, \tilde{v}_{m-2} \text{ are sequentially updated during } (t_L, t_L+k)) = \mathbb{P}(S \leqslant k) = \mathbb{P}(N_k \geqslant m-2).$$

Noting that there are at most $(\Delta k)^m$ cycles of length m containing a_L finishes the proof. ☐

We now prove Lemma 5.8. Fix $\gamma_L \in \Gamma_L$ and $v, v' \in a_L$. We define the events $\mathsf{M}^{\mathrm{d}}_{\ell}$, $\mathsf{M}^{\mathrm{d}}_{L+1}(t')$ and $\mathsf{M}^{\mathrm{d}}_{L+2}(t)$ in a similar fashion as $\mathsf{M}^{\mathrm{d}}_{\ell}$, $\mathsf{M}^{\mathrm{d}}_{L+1}(t)$ in the proof of Lemma 5.6. Observe that any $\gamma_{L+2} \equiv \gamma_{L+2}(v, v', t, t')$, must satisfy that $L+1, L+2 \in I_{\mathrm{d}}$. By Campbell's theorem,

$$\mathbb{E}[N_2^{\mathrm{proj}} \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] = \mathbb{E}\Big[\sum_{t>t':(v,t,1),(v',t',1)\in\xi} \mathbf{1}\{\mathsf{M}^{\mathrm{d}}_{L+1}(t')\} \cdot \mathbf{1}\{\mathsf{M}^{\mathrm{d}}_{L+2}(t)\} \Bigm| \mathsf{N}_L \Big]$$
$$= \frac{1}{4}\int_{t_L}^{\infty}\int_{t_L}^{t} \mathbb{P}(\mathsf{M}^{\mathrm{d}}_{L+2}(t), \mathsf{M}^{\mathrm{d}}_{L+1}(t') \mid \mathsf{N}_L, \mathsf{A}_{L+1}(t'), \mathsf{A}_{L+2}(t)) dt dt'.$$

Proof of Lemma 5.8. Following a similar argument of Lemma 5.10 we can show that for each $\gamma_{L+2}(v, v', t, t') \in \Gamma_{L+2}$,

$$\begin{aligned}
&\mathbb{P}(\mathsf{M}^{\mathrm{d}}_{L+2}(t), \mathsf{M}^{\mathrm{d}}_{L+1}(t') \mid \mathsf{N}_L, \mathsf{A}_{L+1}(t'), \mathsf{A}_{L+2}(t))\\
&\leqslant \mathbb{P}(\mathsf{B}_{L+1}(t'), \mathsf{B}_{L+2}(t)) \cdot \mathbb{P}(\mathsf{D}^2_{L+2}(t) \mid \mathsf{N}^{\circ}_L, \mathsf{C}'_{L+2}(t), \mathsf{B}_{L+1}(t'), \mathsf{B}_{L+2}(t)))\\
&\leqslant 2^{1-k}\mathbb{P}(\mathsf{B}_{L+1}(t'), \mathsf{B}_{L+2}(t)),
\end{aligned} \tag{37}$$

where in the second step we ignored the event $\mathsf{D}^2_{L+1}(t')$ and in the last step we used the independency between $Y_t(a_L)$ and $Y_{t_L}(a_L)$ given $\mathsf{C}'_{L+2}(t)$.

Now let $T_1 \equiv T_+(a_L; a_L, t_L)$ and $T_2 \equiv T_+(a_L; a_L, T_1)$. By monotonicity of deactivation time, event $\mathsf{B}_{L+2}(t) \subseteq \{T_2 > t\}$. Integrating the RHS of (37) over t, t', we have

$$\mathbb{E}[N_2^{\mathrm{proj}} \mid \gamma_L \in \tilde{\Gamma}_{\mathrm{proj},L}] \leqslant 2^{-(k+1)} \int_{t_L}^{\infty} (t - t_L)\mathbb{P}(T_2 > t) dt = 2^{-(k+1)}\mathbb{E}(T_2 - t_L)^2.$$

Both $(T_1 - t_L)$ and $(T_2 - T_1)$ are distributed as the maximum of k i.i.d. $\mathrm{Exp}(1)$ random variables and they are independent with each other. Therefore a very crude bound gives

$$\mathbb{E}(T_2 - t_L)^2 \leqslant 4\mathbb{E}(T_1 - t_L)^2 \leqslant 4k^2\mathbb{E}_{X\sim\mathrm{Exp}(1)}[X^2] \leqslant 8k^2.$$

Plugging the last inequality into (37) concludes the proof. □

6. From sampling to counting

In this section we derive Corollary 1.2 from our main result. The corollary follows from the rapid mixing of the Markov chain and following lemma, which is an analog of [7, Appendix A] and [1, Lemma 5]. Let $\mathcal{G}(k, \Delta)$ be the set of k-uniform hypergraphs of maximal degree Δ.

Lemma 6.1. *Let k and Δ be positive integers and $\mathcal{G} \subseteq \mathcal{G}(k, \Delta)$ be a subset of $\mathcal{G}(k, \Delta)$ that is closed under removal of hyperedges. Suppose that for each Δ, k, there is a polynomial-time algorithm (in n and $1/\epsilon$) that takes a hypergraph $G = (V, F, E) \in \mathcal{G}$ with at most n vertices, a vertex $v \in V$ and an $\epsilon > 0$ and outputs a quantity $p(v; G)$ satisfying*

$$\left| \frac{p(v; G)}{\mathbb{P}_G(\sigma_v = 0)} - 1 \right| < \epsilon.$$

with probability $1 - \epsilon/n$, where σ is a uniformly sampled independent set on G. Then there exists an FPRAS which approximates $Z(G)$ for all hypergraphs in $\mathcal{G}$.

Proof. The proof is a slight modification from the argument in [1, Lemma 5] which we only include here for the sake of completeness. Fix $\epsilon > 0$ and $G = (V, F, E) \in \mathcal{G} \subseteq \mathcal{G}(k, \Delta)$. Without loss of generality, we suppose $V \equiv [n] \equiv \{1, \ldots, n\}$. Let $G_0 \equiv G$ and for each $1 \leqslant i \leqslant n-1$, let G_i be the remaining hypergraph after removing the first i vertices

$[i] \equiv \{1, \ldots, i\}$ along with all hyperedges containing at least one vertex in $[i]$. The set of independent sets on G_i can be naturally identified with the subset of independent sets on G satisfying $\sigma|_{\{1,\ldots,i\}} = \underline{0}$. We have

$$\frac{1}{Z(G)} = \mathbb{P}_G(\underline{\sigma} = \underline{0}) = \mathbb{P}(\sigma_1 = 0) \prod_{i=2}^{n} \mathbb{P}_G(\sigma_i = 0 \mid \sigma_{[i-1]} = 0) = \prod_{i=1}^{n} \mathbb{P}_{G_{i-1}}(\sigma_i = 0).$$

By assumption, the set $\mathcal{G}$ is closed under the removal of hyperedges, thus if $G \in \mathcal{G}$, then so is every G_i, for all $1 \leqslant i \leqslant n-1$. Consequently, we can compute (in $\mathrm{poly}(n, 1/\epsilon)$ time) quantities $p_i \equiv p(i; G_{i-1})$ such that

$$\left| \frac{p(i; G_{i-1})}{\mathbb{P}_{G_{i-1}}(\sigma_i = 0)} - 1 \right| < \frac{\epsilon}{2n},$$

with probability $1 - \epsilon/n$. Letting $\hat{Z}(G) \equiv \prod_{i=1}^{n} p(i; G_{i-1})$ be the output concludes the proof. □

Proof of Corollary 1.2. Following the statement of Lemma 6.1, we set $\mathcal{G} = \mathcal{G}(k, \Delta)$ and describe the $p(v, G)$-outputting algorithm as follows: Given hypergraph G and $n, \epsilon > 0$, let $t_{\mathrm{mix}} = O(n \log n)$ be the mixing time of the Glauber dynamics of hypergraph independent set on G and let N, M be two large integers to be determined shortly. We run the Glauber dynamics N times for $M \cdot t_{\mathrm{mix}}$ steps, starting from the all zeros configuration, and record the configuration at time $M \cdot t_{\mathrm{mix}}$ of the r'th sample by $\sigma^{(r)}$. We set $M \equiv 1 + 2\lceil |\log \epsilon| / \log 2 \rceil$. By the submultiplicity property $t_{\mathrm{mix}}(2^{-i}) \leqslant i t_{\mathrm{mix}}$ [6, page 55] we have that

$$\left|\mathbb{P}(\sigma_v^{(1)} = 0) - \mathbb{P}_G(\sigma_v = 0)\right| \leqslant 2 \|\mathbb{P}(\sigma^{(1)} = \cdot) - \mathbb{P}_G(\sigma = \cdot)\|_{\mathrm{TV}} \leqslant 2^{-(M-1)} \leqslant \epsilon^2 < \epsilon/4,$$

where σ is a uniformly chosen independent set. We set $N \equiv 32\lceil |\log \epsilon| \rceil / \epsilon^2$ and $p(v; G) \equiv \frac{1}{N} \sum_{r=1}^{N} \mathbf{1}\{\sigma_v^{(r)} = 0\}$. By Azuma-Hoeffding's inequality,

$$\mathbb{P}\left(|p(v; G) - \mathbb{P}(\sigma_v^{(1)} = 0)| > \epsilon/4\right) \leqslant e^{-N\epsilon^2/32} \leqslant \epsilon.$$

Note that $\mathbb{P}_G(\sigma_v = 0) \geqslant 1/2$ for any hypergraph $G = (V, F)$ and all $v \in V$. Combining the last two displays then guarantees that $|p(v; G)/\mathbb{P}(\sigma_v = 0) - 1| < \epsilon$ with probability $1 - \epsilon$. The total running time of our algorithm is $N \cdot M \cdot t_{\mathrm{mix}}$, which by Theorem 1.1 is $\mathrm{poly}(n, 1/\epsilon)$. □

References

[1] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Štefankovič. Approximation via correlation decay when strong spatial mixing fails. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 45:1–45:13, 2016.

[2] Magnus Bordewich, Martin Dyer, and Marek Karpinski. Stopping times, metrics and approximate counting. In *International Colloquium on Automata, Languages, and Programming*, pages 108–119. Springer, 2006.

[3] Magnus Bordewich, Martin Dyer, and Marek Karpinski. Path coupling using stopping times and counting independent sets and colorings in hypergraphs. *Random Structures & Algorithms*, 32(3):375–399, 2008.

[4] Heng Guo, Mark Jerrum, and Jingcheng Liu. Uniform sampling through the Lovász Local Lemma. *arXiv preprint arXiv:1611.01647*, 2016.

[5] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on computing*, 22(5):1087–1116, 1993.

[6] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.

[7] Jingcheng Liu and Pinyan Lu. FPTAS for counting monotone CNF. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1531–1548. SIAM, 2015.

[8] Eyal Lubetzky and Allan Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Mathematical Journal*, 153(3):475–510, 2010.

[9] Eyal Lubetzky and Allan Sly. Information percolation and cutoff for the stochastic Ising model. *Journal of the American Mathematical Society*, 2015.

[10] Ankur Moitra. Approximate counting, the Lovasz Local Lemma and inference in graphical models. *arXiv preprint arXiv:1610.04317*, 2016.

[11] Robin A Moser. A constructive proof of the lovász local lemma. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 343–350. ACM, 2009.

[12] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *Journal of Statistical Physics*, 155(4):666–686, 2014.

[13] Allan Sly. Computational transition at the uniqueness threshold. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 287–296. IEEE, 2010.

[14] Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 361–369. IEEE, 2012.

[15] Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 140–149. ACM, 2006.